



CENTRO INTERNACIONAL DE ESTUDOS
DE DOUTORAMENTO E AVANZADOS
DA USC (CIEDUS)

TESIS DE DOCTORADO

VALIDACIÓN DE CONSTRUCTO Y DE CRITERIO DE LA REFLEXIBILIDAD COGNITIVA

Inmaculada Otero Moral

ESCUELA DE DOCTORADO INTERNACIONAL EN ARTES Y HUMANIDADES,
CIENCIAS SOCIALES Y JURÍDICAS

PROGRAMA DE DOCTORADO EN PSICOLOGÍA DEL TRABAJO Y LAS
ORGANIZACIONES, JURÍDICA-FORENSE Y DEL CONSUMIDOR Y USUARIO

SANTIAGO DE COMPOSTELA
2019





DECLARACIÓN DE LA AUTORA DE LA TESIS

Validación de Constructo y de Criterio de la Reflexibilidad Cognitiva

Dña. Inmaculada Otero Moral

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento, y declaro que:

- 1) La tesis abarca los resultados de la elaboración de mi trabajo.
- 2) De ser el caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.
- 3) La tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.
- 4) Confirmando que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.

En Santiago de Compostela, a 9 de octubre de 2019

Fdo. Inmaculada Otero Moral





AUTORIZACIÓN DE LOS DIRECTORES DE LA TESIS

Validación de Constructo y de Criterio
de la Reflexibilidad Cognitiva

D. Jesús F. Salgado Velo
Dña. Silvia Moscoso Ruibal

INFORMAN:

*Que la presente tesis, se corresponde con el trabajo realizado por Dña. **Inmaculada Otero Moral**, bajo nuestra dirección, y autorizamos su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC y que, como directores de ésta, no incurrir en las causas de abstención establecidas en la Ley 40/2015.*

En Santiago de Compostela, a 9 de octubre de 2019

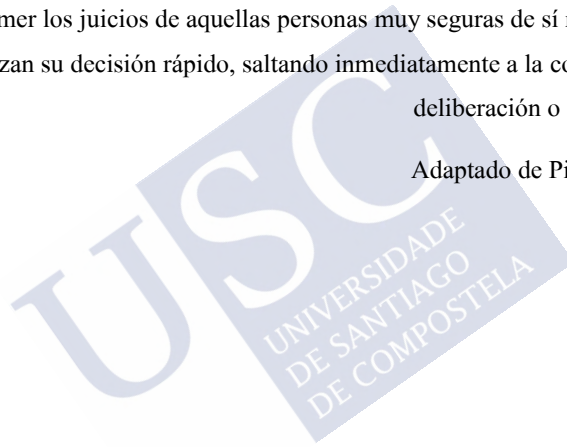
Fdo. D. Jesús F. Salgado Velo

Fdo. Dña. Silvia Moscoso Ruibal



“Deberíamos temer los juicios de aquellas personas muy seguras de sí mismas, aquellas que alcanzan su decisión rápido, saltando inmediatamente a la conclusión sin una deliberación o arrepentimiento”

Adaptado de Piero Calamandrei





*A mi abuelo Pepe,
por impulsarme siempre a crecer*





AGRADECIMIENTOS

La realización de una tesis doctoral es el resultado de mucho esfuerzo y de largas horas frente al ordenador, pero, sobre todo, es el resultado de una gran constancia y perseverancia por alcanzar ese logro. El mérito, sin embargo, no puede ser atribuible únicamente al doctorando ya que son muchas las personas que, de un modo u otro, contribuyen en el proceso y conclusión del trabajo. A través de estas líneas deseo mostrar mi más sincero agradecimiento a todas y cada una de ellas.

En primer lugar, quiero agradecer a mis directores de tesis, el Profesor Jesús F. Salgado y la Profesora Silvia Moscoso, la confianza depositada al darme la oportunidad de formar parte de su grupo de investigación. Todos los conocimientos, la formación, experiencia y calidad humana que me he topado a vuestro lado hacen que estos años sean una de las etapas más memorables de mi vida. Asimismo, quiero agradecerles la dirección y seguimiento de esta tesis. Si hoy en día es lo que es, indudablemente es gracias a ellos. Al Profesor Jesús F. Salgado quiero agradecerle, particularmente, su inestimable ayuda con los datos, donde sus conocimientos metodológicos y estadísticos han contribuido a la calidad de este trabajo. También quiero agradecerle todo el apoyo y las muestras de cariño que me ha dado a lo largo de los años y, especialmente, en los últimos meses de trabajo. Del mismo modo, me gustaría agradecerle la dedicación y contribución realizada a mi formación como científica y profesional. El entusiasmo, el amor por el trabajo y, especialmente, el interés por los números y las matemáticas son valores que, en parte, he adquirido a su lado. A la profesora Silvia Moscoso quiero agradecerle, particularmente, las constantes revisiones del texto y todas y cada una de sus sugerencias. Gracias a ello, estoy convencida de que su lectura, comprensión y seguimiento se hace de forma más fácil y entretenida. Asimismo, quiero mostrarle mi gratitud por las palabras de serenidad y todos los gestos de cariño que ha tenido conmigo en los momentos de mayor dificultad. Finalmente, también quiero agradecerle su apoyo y dedicación para ayudarme a crecer como investigadora.

Debo agradecer, también, a la Xunta de Galicia y al Ministerio de Educación, Cultura y Deporte por la ayuda económica otorgada para realizar la tesis y la estancia doctoral. Asimismo, debo agradecer al Ministerio de Economía y Competitividad que, a través de los proyectos PSI2014-56615-P y PSI2017-87603-P, concedidos al Profesor Jesús F. Salgado, han financiado la recogida de datos de este trabajo y han permitido mi continuación como científica. También agradezco a la Facultad de Relaciones Laborales, a la Universidad de Santiago de Compostela y a las personas que allí trabajan por facilitarme el acceso a las instalaciones y a los recursos pertinentes para alcanzar mi propósito. Gracias, también, a Loli Grela, María Jesús Souto, Mario Lado, Joaquín Vidal y a todas aquellas personas que, con su interés por mi trabajo y mi bienestar, me han hecho sentir querida en la facultad.

Mi agradecimiento también al Profesor Ramón Rico y a la Bussiness School de la Universidad de Western of Australia (Perth) por haberme permitido realizar mi estancia doctoral con ellos y haberme acogido con tanto cariño. Gracias también a Kiri Adams por mostrarme su trabajo en el laboratorio y por involucrarme en el proyecto. Del mismo modo, quiero resaltar la calidez de todas las personas que he conocido allí y de las que guardo un bonito recuerdo: Roberto, Minou, Chloe, Justin, Estelle y Calle. Pero, en especial, quiero agradecer a ti, Rina, a mi “mamá australiana”, que desde un principio me has acogido como a una hija y solo me has colmado de cariño y atenciones. Sin duda, haber realizado la estancia doctoral con todos vosotros ha sido una de mis mejores experiencias académicas y personales.

Quiero mostrar también mi gratitud a mis compañeras de trabajo, de despacho y, por qué no, de alguna que otra juerga, porque su apoyo y colaboración han sido relevantes en este trabajo. Gracias a Dámaris Cuadrado por su ayuda durante la fase de codificación de los datos, por las muestras de ánimo y por los consejos de trabajo. Gracias a Alexandra Martínez porque, como “compañera de batalla”, hemos podido compartir esta experiencia, acompañándonos, entendiéndonos y consolándonos mutuamente. Quiero agradecer, también, a Pamela Alonso por sus incansables revisiones, por todas y cada una de las palabras de apoyo y elogio a mi trabajo y por ser, en fin, tal y como es. Del mismo modo, quiero agradecer a mi compañero José Carlos Tavares por todo el apoyo brindado, por sus ilustrativas charlas y por no olvidarse nunca de venir a visitarme al laboratorio.

Pero, sobre todo, debo mostrar mi gratitud a toda mi familia y amigos, por su incondicional apoyo, por estar ahí siempre, a mi lado y por demostrarme su inmenso cariño de tan variadas formas. En especial, quiero resaltar mi más inmenso agradecimiento a mis adorados Susa y Manolo, porque si ellos no me hubiesen clavado la espinita de la investigación, nunca me habría atrevido a dar este paso. Gracias por las constantes muestras de cariño, por todo el apoyo brindado y por la ilusión y admiración que me habéis transmitido con cada palabra. En efecto, una parte de esta tesis es un mérito exclusivamente vuestro. Quiero agradecer también a mis padres, Mari y Manolo, por todo el esfuerzo y trabajo para poder pagarnos unos estudios a mi hermana y a mí. Este es el mejor legado que unos padres le pueden dejar a sus hijos. Gracias también a mi madre por inculcarme el valor del trabajo y estar día tras día animándome a seguir estudiando. Sin esa constancia tuya, posiblemente, no habría llegado hasta aquí. Gracias a mi hermana Cristina, a mi cuñado David y a mis sobrinos, Yara y Mateo, por apoyarme en todo el proceso, aun cuando no entendíais la necesidad de dedicarle tantas horas a este trabajo. Gracias, especialmente, a mi ahijado Mateo por todas tus ocurrencias y travesuras y porque, sin ser consciente de ello, has llenado mi vida de luz y alegría. Gracias también a mis tíos, Higinio y Ana, y a mis primos, Mireya, Christian, Ely y David, por interesaros por mi trabajo y alegraros de mis logros. Gracias a Ely por su inestimable ayuda en el cierre de la búsqueda bibliográfica y por ser conmigo como una hermana.

Finalmente, quiero agradecer a mis amigos Diego, Paula, Marina, Ramón, Andrea, Isa, Dani y Jose, por haber compartido conmigo tantos años de amistad, por haberme aguantado en mis peores días y por haberme apoyado cuando más lo he necesitado. Gracias también por haber compartido conmigo vuestros buenos momentos y por tantos y tan bellos recuerdos. Solo puedo desear que sigamos así de unidos por mucho tiempo más.





RESUMEN

Esta tesis doctoral se ha llevado a cabo con el propósito de profundizar en el conocimiento de la flexibilidad cognitiva (RC). El desarrollo de este concepto parte de los Modelos Duales de Procesamiento donde se defiende que el razonamiento humano está guiado por dos tipos de procesamiento cognitivos. Uno automático e intuitivo, que opera de forma rápida y con poco o ningún esfuerzo y otro analítico y reflexivo, que opera de forma más lenta y con esfuerzo y concentración. Para evaluar este constructo, Frederick (2005; Kahneman y Frederick, 2002) desarrolló el Test de Reflexibilidad Cognitiva (TRC-3). Desde entonces, autores de muy diversas disciplinas se han interesado en validar este instrumento. Sin embargo, apenas se ha investigado en el ámbito de la Psicología del Trabajo. El principal objetivo de esta tesis es estudiar la medida y el concepto de RC en este ámbito. Con esta finalidad, se presentan cuatro estudios empíricos.

En el estudio 1 se analizan las propiedades psicométricas de una nueva medida de RC (TRC-13; Salgado, 2014a) y se comparan los resultados con los encontrados en el test original (TRC-3; Frederick, 2005). Los datos muestran que el nuevo test es un instrumento adecuado de evaluación de la RC y que presenta mejores propiedades psicométricas que el TRC-3. También se sugiere que los hombres tienden a alcanzar puntuaciones significativamente más altas que las mujeres.

En el estudio 2 se exploran los efectos del orden de presentación de los ítems del test sobre el resultado y el tiempo de respuesta al TRC-13. Los resultados concluyen que colocando los 3 ítems originales (Frederick, 2005) en medio de los nuevos (Salgado, 2014a) se pueden mejorar los resultados en el TRC-13 y en los 3 ítems originales. También se muestra que el orden de presentación de los ítems no afecta al tiempo que se invierte en responder al TRC-13, aunque sí incrementa significativamente el tiempo de respuesta a los ítems originales o nuevos cuando estos se posicionan al inicio del test.

En el estudio 3 se examina la validez de constructo y de criterio del TRC. Los resultados muestran que la RC es un constructo principalmente cognitivo, aunque también se relaciona con variables de personalidad (disposiciones de pensamiento y factores de personalidad). Asimismo, sugieren que las variables que determinan la RC en hombres y mujeres podrían ser diferentes. Finalmente, se muestra que la RC es un predictor válido de diversos criterios académicos y ocupacionales y que contribuye en la predicción múltiple añadiendo validez sobre otras variables.

En el cuarto estudio se presentan 3 meta-análisis sobre la validez de constructo y de criterio del TRC. En el primero de ellos, se confirma la relación entre la capacidad cognitiva y la RC. En el segundo, se estima la validez teórica y operativa y la generalización de su validez para predecir el desempeño académico. En el tercer meta-análisis se determina el tamaño del efecto teórico de las diferencias en función del sexo y se confirma que los hombres alcanzan mejores resultados en el TRC que las mujeres.

Finalmente, se discuten las implicaciones teóricas y prácticas de los resultados y se sugieren futuras líneas de investigación.

Palabras clave: *flexibilidad cognitiva; capacidad mental general; personalidad; desempeño académico y ocupacional; validez de constructo y de criterio*



RESUMO

Esta tese de doutoramento realizouse co propósito de afondar no coñecemento da flexibilidade cognitiva (RC). Este concepto parte dos Modelos Duais de Procesamento, onde se defende que o razoamento humano está guiado por dous tipos de procesamentos cognitivos. Un automático e intuitivo, que opera de forma rápida e con pouco ou ningún esforzo e outro analítico e reflexivo, que opera de forma máis lenta e con esforzo e concentración. Para avaliar este construto, Frederick (2005; Kahneman y Frederick, 2002) propuxo o Test de Flexibilidade Cognitiva (TRC-3). Dende entón, autores de moi diversas disciplinas interesáronse en validar este instrumento. Sen embargo, apenas se investigou no ámbito da Psicoloxía do Traballo. O principal obxectivo de esta tese é estudar a medida e o concepto de RC neste ámbito. Con esta finalidade, preséntanse catro estudos empíricos.

No estudo 1 analízanse as propiedades psicométricas dunha nova medida de RC (TRC-13; Salgado, 2014a) e compáranse os resultados cos atopados no test orixinal (TRC-3; Frederick, 2005). Os datos amosan que o novo test é un instrumento adecuado de avaliación da RC e presenta mellores propiedades psicométricas que o TRC-3. Tamén se indica que os homes tenden a alcanzar puntuacións significativamente máis altas que as mulleres.

No estudo 2 explóranse os efectos da orde de presentación dos ítems no test sobre o resultado e o tempo de resposta ao TRC-13. Os resultados conclúen que colocando os 3 ítems orixinais (Frederick, 2005) no medio dos novos (Salgado, 2014a) pódense mellorar os resultados no TRC-13 e nos 3 ítems orixinais. Tamén se amosa que a orde de presentación dos ítems non afecta ao tempo que se inviste en responder ao TRC-13, aínda que si incrementa significativamente o tempo de resposta aos ítems orixinais ou novos cando estes se dispoñen ao inicio do test.

No estudo 3 examínase a validez de construto e de criterio do TRC. Os resultados indican que a RC é un construto principalmente cognitivo, aínda que tamén se relaciona con variables de personalidade (disposicións de pensamento e factores de personalidade). Do mesmo xeito, propónse que as variables que determinan a RC nos homes y nas mulleres poderían ser diferentes. Finalmente, amósase que a RC é un preditor válido de diversos criterios académicos e ocupacionais e que contribúen na predición múltiple engadindo validez sobre outras variables.

No cuarto estudo preséntanse 3 meta-análises sobre a validez de construto e de criterio do TRC. No primeiro, confírmase a relación entre a capacidade cognitiva e a RC. No segundo, estímase a validez teórica e operativa e a xeneralización da súa validez para predicir o desempeño académico. No terceiro meta-análise determínase o tamaño do efecto teórico das diferenzas en función do sexo e confírmase que os homes alcanzan mellores resultados no TRC que as mulleres.

Finalmente, discútense as implicacións teóricas e prácticas dos resultados e propóñense futuras liñas de investigación.

Palabras chave: *flexibilidade cognitiva; capacidade mental xeral; personalidade; desempeño académico e ocupacional; validez de construto e de criterio*



ABSTRACT

This PhD dissertation has been carried out with the purpose of better knowing the construct of cognitive reflection (CR). This concept has been developed from Dual-Process Models, which postulate that human reasoning is controlled by two types of cognitive processes. One of them is automatic, intuitive, and operates fast with little or any effort. The second one is analytic, reflexive, and operates faster than the first, with effort and concentration. To evaluate the construct, Frederick (2005; Kahneman & Frederick, 2002) developed the Cognitive Reflection Test (TRC-3). Since that moment, authors from a variety of disciplines have been interested in validating this instrument. However, it has not been yet studied in the Industrial Psychology field. The main goal of this dissertation is to study the CR measure and the CR in this domain. For this purpose, four empirical studies were carried out.

Study 1 analyses the psychometric properties of a new CR measure (CRT-13; Salgado, 2014a) and it compares them with the evidence found for the original test (CRT-3; Frederick, 2005). Results show that the new test is a good instrument to assess CR and that it has better psychometric properties than the CRT-3. The study also suggests that men are more likely to achieve significantly higher scores than women.

Study 2 explores the effects that the order of the items has on the TRC-13 performance and on time reaction. Results suggest that the performance on the TRC-13 (Salgado, 2014a) and on the original items (Frederick, 2005) can be improved if the three original items are placed in the middle of the new ones. Moreover, the order of the items does not affect the TRC-13 time reaction. However, the original or the new items time reaction could increase when they are placed at the beginning of the test.

Study 3 examines the construct validity and the criterion validity of the CRT. Results show that CR is essentially a cognitive construct, although it is also related to personality variables (thinking dispositions and personality dimensions). Furthermore, they suggest that the variables that determine CR in men and women could be different. Finally, they show that CR is a valid predictor of several academic and occupational criteria and that it produces incremental validity over the other variables.

Study 4 presents three meta-analyses on the construct validity and criteria validity of the CRT. The first meta-analysis shows a relationship between cognitive ability and CR. The second one estimates the true validity and the operational validity of CR with regards to academic performance. CR also showed validity generalization to predict this criterion. The third meta-analysis determines the true effect size of the sex differences in CR and it confirms that men achieve higher scores than women in the CRT.

Theoretical and practical implications of the results are finally discussed. Suggestions for future research are also stated.

Keywords: *cognitive reflection; general mental ability; personality; academic performance and occupational performance; construct validity and predictive validity*



ÍNDICE

INTRODUCCIÓN.....	33
-------------------	----

MARCO CONCEPTUAL

CAPÍTULO 1. Modelos de Procesamiento Dual (<i>Dual-Process Models</i>)	41
1. INTRODUCCIÓN	41
2. MODELOS DE PROCESAMIENTO DUAL (MPD)	42
2.1. <i>Sistemas Asociados a los Procesamientos</i>	44
2.2. <i>Memorias Asociadas a los Procesamientos</i>	52
2.3. <i>Categorías y Atributos de los dos Tipos de Procesamiento</i>	55
2.3.1. Características de los Procesamientos	55
2.3.2. Elementos, Información y Principios Mediante los que Operan	61
2.3.3. Funciones de los Procesamientos	65
2.3.4. Arquitectura de los Procesamientos	67
2.3.5. Modulación de la Activación de los Procesamientos	70
2.3.5.1. Variables Motivacionales	70
2.3.5.2. Diferencias Individuales	71
2.3.5.3. Variables Situacionales	73
2.3.6. Evolución Biológica	75
3. OTROS MODELOS COGNITIVOS	75
3.1. <i>Modelo de Proceso Único</i>	75
3.2. <i>Modelos Tricotómicos</i>	76
3.2.1. Teoría General de la Racionalidad (OSCAR: General Theory of Rationality) de Pollock (1989, 1991)	76
3.2.2. Modelo Híbrido de Procesamiento de Evans (2009)	76
3.2.3. Modelo Tripartito de Procesamiento Cognitivo de Stanovich (2011)	77
3.3. <i>Modelos de Múltiples Procesos</i>	77
3.3.1. Modelos que Definen el Procesamiento como un Continuo	77
3.3.2. Modelos que Postulan una Multiplicidad de Procesos	78
3.3.2.1. Modelo de Glöckner y Witteman (2010)	78

3.3.2.2. Modelo de Conrey, Sherman, Gawronski, Hugenberg y Groom (2005)	79
3.3.2.3. Otros modelos.....	79
4. CRÍTICAS A LOS MODELOS DUALES DE PROCESAMIENTO	80
CAPÍTULO 2. Reflexibilidad Cognitiva (RC) y Test de Reflexibilidad Cognitiva (TRC)	83
1. INTRODUCCIÓN	83
2. CONCEPTO DE REFLEXIBILIDAD COGNITIVA (RC) Y TEST DE REFLEXIBILIDAD COGNITIVA (TRC)	84
3. TEST DE REFLEXIBILIDAD COGNITIVA (TRC)	85
3.1. <i>Procesamiento Cognitivo para Responder al TRC</i>	89
3.2. <i>Metacognición del Procedimiento de Respuesta a los Ítems</i>	93
3.2.1. Experiencia Metacognitiva del Sentimiento de Respuesta Correcta y de la Percepción del Desempeño de uno Mismo en el TRC	97
3.2.2. Experiencia Metacognitiva de la Percepción del Desempeño de Otros en el TRC.....	99
3.3. <i>Tiempo de Reacción o Fluidez de la Respuesta</i>	102
3.4. <i>Otros Determinantes del Desempeño en el TRC</i>	103
3.4.1. Relacionados con la Tarea.....	103
3.4.1.1. Redacción de las Instrucciones.....	103
3.4.1.2. Fluidez de la Fuente de los Ítems o Fluidez Perceptual	104
3.4.1.3. Dificultad del Idioma.....	106
3.4.1.4. Secuencia/Orden de los Ítems	106
3.4.2. Relacionados con el Contexto	107
3.4.2.1. Posición del Test Dentro del Estudio Experimental.....	107
3.4.2.2. Formato de Presentación del Test (Lápiz y Papel vs. Ordenador).....	107
3.4.2.3. Tarea Incentivada	107
3.4.2.4. Feedback.....	108
3.4.2.5. Momento del Día (Acrofase).....	108
3.4.2.6. Carga Cognitiva.....	109
3.4.2.7. Agotamiento del Autocontrol.....	109
3.4.2.8. Estrés	110
3.4.2.9. Efecto Priming Previo	110
3.5. <i>Limitaciones del TRC</i>	111

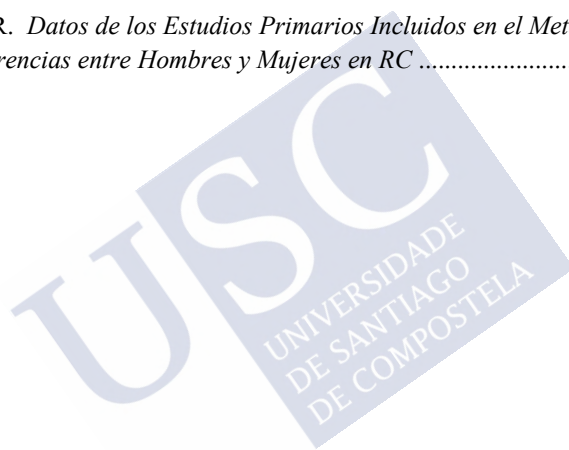
CAPÍTULO 3. Evidencias de Validez de Constructo y Criterio de la Medida de Reflexibilidad Cognitiva (RC)	117
1. INTRODUCCIÓN	117
2. EVIDENCIAS DE VALIDEZ DE CONSTRUCTO	117
2.1. ¿Es el TRC una Medida de Capacidad Cognitiva?	118
2.2. ¿Es el TRC una Medida de Capacidad Matemática y Pensamiento Racional?	125
2.3. ¿Es el TRC una Medida de Capacidad Cognitiva, Pensamiento Racional y Disposición hacia un Pensamiento de Mente Abierta?	138
3. EVIDENCIAS DE VALIDEZ BASADAS EN LA RELACIÓN CON OTRAS VARIABLES	145
3.1. Personalidad	145
3.1.1. Modelo de los Big Five Factors (BFF)	145
3.1.2. Conductas Deshonestas	147
3.2. Variables Demográficas	148
3.2.1. Sexo	148
3.2.2. Edad	152
4. EVIDENCIAS DE VALIDEZ DE CRITERIO	153
4.1. Desempeño Académico y Reflexibilidad Cognitiva	154
4.2. Desempeño Ocupacional y Reflexibilidad Cognitiva	155

PARTE EMPIRICA

ESTUDIO 1. Propiedades Psicométricas de una Versión Nueva del TRC	159
1. INTRODUCCIÓN	159
2. MÉTODO	160
3. RESULTADOS	167
4. DISCUSIÓN	185
ESTUDIO 2. Efectos del Orden de Presentación de los Ítems sobre el Resultado en el TRC	189
1. INTRODUCCIÓN	189
2. MÉTODO	190
3. RESULTADOS	193
4. DISCUSIÓN	203

ESTUDIO 3. Validez de Constructo y de Criterio del Test de Reflexibilidad	
Cognitiva	207
1. INTRODUCCIÓN	207
2. MÉTODO.....	212
3. RESULTADOS.....	228
4. DISCUSIÓN	255
ESTUDIO 4. Meta-Análisis de la Relación de la RC, Capacidad Mental	
General y Desempeño Académico	259
1. INTRODUCCIÓN	259
2. PROCEDIMIENTO	261
3. MÉTODO DE META-ANÁLISIS	269
4. RESULTADOS.....	275
CONCLUSIONES GENERALES	287
REFERENCIAS	295
APÉNDICES	325
APÉNDICE A. <i>Distribución e Integración de las Puntuaciones del TRC</i> <i>Original</i>	327
APÉNDICE B. <i>Nuevos Ítems de RC Desarrollados/Adaptados por otros Autores</i>	339
APÉNDICE C. <i>Distribución de las Puntuaciones en las Nuevas Versiones del</i> <i>TRC</i>	343
APÉNDICE D. <i>TRC-13 (Salgado, 2014a) y Estadísticos Descriptivos del Test</i>	347
APÉNDICE E. <i>Medias, Desviaciones Típicas y Diferencias en el Desempeño en</i> <i>el TRC según la Condición Experimental (Hombres)</i>	351
APÉNDICE F. <i>Medias, Desviaciones Típicas y Diferencias en el Desempeño en</i> <i>el TRC según la Condición Experimental (Mujeres)</i>	353
APÉNDICE G. <i>Medias y Desviaciones Típicas de los Tiempos de Respuesta (en</i> <i>Segundos) del TRC según la Condición Experimental (Hombres)</i>	355
APÉNDICE H. <i>Medias y Desviaciones Típicas de los Tiempos de Respuesta (en</i> <i>Segundos) del TRC según la Condición Experimental (Mujeres)</i>	357
APÉNDICE I. <i>Medias, Desviaciones Típicas y Correlaciones Observadas entre</i> <i>las Variables (Muestra Total)</i>	358
APÉNDICE J. <i>Medias, Desviaciones Típicas y Correlaciones Observadas entre</i> <i>las Variables (Hombres)</i>	360
APÉNDICE K. <i>Medias, Desviaciones Típicas y Correlaciones Observadas</i> <i>entre las Variables (Mujeres)</i>	362

APÉNDICE L. <i>Efectos Directos en la Predicción de la RC (Muestra Total)</i>	365
APÉNDICE M. <i>Efectos Indirectos en la Predicción de la RC (Muestra Total)</i>	367
APÉNDICE N. <i>Predicción Múltiple del Desempeño en el TRC-13 en la Muestra de Hombres y Mujeres</i>	369
APÉNDICE Ñ. <i>Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Relación de la CMG y la RC Excluyendo las Numeracy Scales</i>	385
APÉNDICE O. <i>Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Relación de la CMG y la RC Incluyendo las Numeracy Scales</i>	387
APÉNDICE P. <i>Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Relación entre las Capacidades Cognitivas Específicas y la RC</i>	389
APÉNDICE Q. <i>Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Validez de la RC para Predecir el Desempeño Académico</i>	391
APÉNDICE R. <i>Datos de los Estudios Primarios Incluidos en el Meta-Análisis sobre Diferencias entre Hombres y Mujeres en RC</i>	393





ÍNDICE DE TABLAS

Tabla 1. <i>Lista no Exhaustiva de los Principales Modelos de Procesamiento y Nomenclatura de los Procesamientos o Sistemas Propuestos</i>	45
Tabla 2. <i>Categorías y Atributos de los dos Tipos de Procesamientos</i>	57
Tabla 3. <i>Clasificación de los MPD de Acuerdo con la Arquitectura Propuesta por Elqayam (2009) y Evans (2008)</i>	69
Tabla 4. <i>Test de Reflexibilidad Cognitiva (TRC)</i>	86
Tabla 5. <i>Sesgos Cognitivos, Definición y Relación con el TRC</i>	131
Tabla 6. <i>Características de la Muestra (N = 1,367)</i>	161
Tabla 7. <i>Ítems del TRC Desarrollado por Salgado (2014a)</i>	163
Tabla 8. <i>Estadísticos Descriptivos del TRC</i>	168
Tabla 9. <i>Frecuencia de las Puntuaciones del TRC-3 (Muestra Total)</i>	169
Tabla 10. <i>Frecuencia de las Puntuaciones del TRC-11 (Muestra Total)</i>	170
Tabla 11. <i>Frecuencia de las Puntuaciones del TRC-14 (Muestra Total)</i>	170
Tabla 12. <i>Porcentajes de Respuestas Correctas y Diferencias en Función del Sexo</i> .	171
Tabla 13. <i>Frecuencia de las Puntuaciones del TRC-3 (Hombres)</i>	173
Tabla 14. <i>Frecuencia de las Puntuaciones del TRC-3 (Mujeres)</i>	173
Tabla 15. <i>Frecuencia de las Puntuaciones del TRC-11 (Hombres)</i>	174
Tabla 16. <i>Frecuencia de las Puntuaciones del TRC-11 (Mujeres)</i>	174
Tabla 17. <i>Frecuencia de las Puntuaciones del TRC-14 (Hombres)</i>	175
Tabla 18. <i>Frecuencia de las Puntuaciones del TRC-14 (Mujeres)</i>	175
Tabla 19. <i>Medias, Desviaciones Típicas y Diferencias en Función del Formato de Administración del TRC</i>	177
Tabla 20. <i>Correlaciones Tetracóricas entre Ítems</i>	179
Tabla 21. <i>Adecuación de la Matriz de Correlaciones</i>	180
Tabla 22. <i>Autovalores y Varianza Explicada Basada en los Autovalores</i>	180
Tabla 23. <i>Cargas Factoriales y Comunalidad de los Ítems</i>	180
Tabla 24. <i>Autovalores y Varianza Explicada Basada en los Autovalores Excluyendo los ítems 8 y 10 o el Ítem 10</i>	181
Tabla 25. <i>Cargas Factoriales y Comunalidad de los Ítems Excluyendo los Ítems 8 y 10 o el Ítem 10</i>	182
Tabla 26. <i>Índices de Ajuste del Modelo Excluyendo el Ítem 10 o los Ítems 8 y 10</i>	182
Tabla 27. <i>Correlaciones entre el TRC-3, TRC-10 y TRC-13</i>	183
Tabla 28. <i>Coeficientes de Fiabilidad de los Test de RC</i>	184
Tabla 29. <i>Características de la Muestra (N = 602)</i>	191
Tabla 30. <i>Distribución de la Muestra por Condiciones Experimentales</i>	193

Tabla 31. <i>Medias, Desviaciones Típicas y Diferencias en el Desempeño en el TRC según la Condición Experimental (Muestra Total)</i>	194
Tabla 32. <i>Diferencias entre Hombres y Mujeres en el Desempeño en el TRC según la Condición Experimental</i>	197
Tabla 33. <i>Medias y Desviaciones Típicas de los Tiempos de Respuesta (en Segundos) del TRC según la Condición Experimental (Muestra Total)</i>	198
Tabla 34. <i>Diferencias entre Hombres y Mujeres en el Tiempo de Respuesta Invertido en el TRC según la Condición Experimental</i>	200
Tabla 35. <i>Coefficientes de Fiabilidad del Tiempo de Respuesta en el TRC-13, TRC-10 y TRC-3</i>	201
Tabla 36. <i>Correlaciones Policóricas entre los TRC y el TR</i>	202
Tabla 37. <i>Características de la Muestra (N = 1,367)</i>	213
Tabla 38. <i>Coefficientes de Restricción en el Rango del TRC en Función de las Variables Criterio y las Muestras de Estudio</i>	214
Tabla 39. <i>Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango del WPT</i>	215
Tabla 40. <i>Coefficientes de Fiabilidad y Coeficiente de Restricción en el Rango del Factor G de Cattell</i>	216
Tabla 41. <i>Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango del RLS</i>	217
Tabla 42. <i>Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango de los Cinco Factores de Personalidad del Inventario IP/5F</i>	218
Tabla 43. <i>Coefficientes de Fiabilidad de los Cinco Factores de Personalidad de la Escala QI5F-Tri</i>	219
Tabla 44. <i>Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango de la Escala Breve de Autocontrol</i>	220
Tabla 45. <i>Coefficientes de Fiabilidad de la Escala NFC</i>	221
Tabla 46. <i>Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango del Test de Figuras Enmascaradas</i>	222
Tabla 47. <i>Coefficientes de Fiabilidad de la Escala de Conductas Personales (CP5)</i>	223
Tabla 48. <i>Coefficientes de Fiabilidad de la Escala CDTE</i>	225
Tabla 49. <i>Coefficientes de Fiabilidad de la Escala CDCE</i>	226
Tabla 50. <i>Coefficientes de Fiabilidad de la Escala CDAN</i>	227
Tabla 51. <i>Coefficientes de Fiabilidad del Test 2 de la Batería GCT</i>	228
Tabla 52. <i>Correlaciones Corregidas entre el TRC y las Medidas de CMG</i>	231
Tabla 53. <i>Correlaciones Corregidas entre el TRC y las Medidas de Personalidad</i> ...	233
Tabla 54. <i>Correlaciones Corregidas entre el TRC y las Medidas de Disposición de Pensamiento</i>	237
Tabla 55. <i>Determinantes de las Disposiciones de Pensamiento (Muestra Total)</i>	240
Tabla 56. <i>Matriz de Correlaciones Corregidas Utilizada en el Análisis de Regresión para Determinar el Desempeño del TRC-13 (Muestra Total)</i> ...	242
Tabla 57. <i>Análisis de Regresión Múltiple para Determinar el Desempeño del TRC-13 (Muestra Total)</i>	243
Tabla 58. <i>Índices de Ajuste del Modelo de Predicción de la RC (Muestra Total)</i>	243

Tabla 59. <i>Correlaciones Corregidas entre el TRC y las Medidas de Desempeño Académico y Ocupacional</i>	246
Tabla 60. <i>Matriz de Correlaciones Corregidas Utilizada en los Análisis de Regresión para Predecir las Medidas Criterio</i>	250
Tabla 61. <i>Análisis de Regresión para Predecir el Desempeño Académico (Notas de Bachillerato)</i>	251
Tabla 62. <i>Análisis de Regresión para Predecir el Desempeño Académico (CAU)</i>	252
Tabla 63. <i>Análisis de Regresión para Predecir el Desempeño Académico (GPA)</i>	253
Tabla 64. <i>Análisis de Regresión para Predecir el Desempeño Ocupacional (GCT)</i> ..	254
Tabla 65. <i>Estudios Integrados que han Empleado Nuevas Versiones del TRC</i>	262
Tabla 66. <i>Instrumentos de Inteligencia Utilizados en los Estudios Primarios</i>	263
Tabla 67. <i>Distribución de los Coeficientes de Fiabilidad de las Variables de Capacidad Cognitiva</i>	271
Tabla 68. <i>Estimación del Coeficiente de Fiabilidad CES de las Nuevas Versiones del TRC según el Número de Ítems</i>	273
Tabla 69. <i>Distribución de los Coeficientes de Fiabilidad de la RC</i>	273
Tabla 70. <i>Distribución de los Coeficientes de Restricción en el Rango (u) de las Variables</i>	275
Tabla 71. <i>Resultados de los Meta-Análisis de la Relación de la Capacidad Cognitiva y la RC, Excluyendo las Numeracy Scales</i>	278
Tabla 72. <i>Resultados de los Meta-Análisis de la Relación de la Capacidad Cognitiva y la RC, Incluyendo las Numeracy Scales</i>	279
Tabla 73. <i>Resultados de los Meta-Análisis de Validez de la RC para Predecir el Desempeño Académico</i>	282
Tabla 74. <i>Resultados del Meta-Análisis de Diferencias en Función del Sexo en la RC</i>	284



ÍNDICE DE FIGURAS

Figura 1. <i>Distribución de las Puntuaciones del TRC-3 (Muestra Total)</i>	169
Figura 2. <i>Distribución de las Puntuaciones del TRC-11 (Muestra Total)</i>	170
Figura 3. <i>Distribución de las Puntuaciones del TRC-14 (Muestra Total)</i>	170
Figura 4. <i>Distribución de las Puntuaciones del TRC-3 (Hombres)</i>	173
Figura 5. <i>Distribución de las Puntuaciones del TRC-3 (Mujeres)</i>	173
Figura 6. <i>Distribución de las Puntuaciones del TRC-11 (Hombres)</i>	174
Figura 7. <i>Distribución de las Puntuaciones del TRC-11 (Mujeres)</i>	174
Figura 8. <i>Distribución de las Puntuaciones del TRC-14 (Hombres)</i>	175
Figura 9. <i>Distribución de las Puntuaciones del TRC-14 (Mujeres)</i>	175
Figura 10. <i>Representación Gráfica del Modelo de Predicción de la RC (Muestra Total)</i>	244



INTRODUCCIÓN

En 2005, Shane Frederick introdujo el concepto de flexibilidad cognitiva (RC) para referirse a la “capacidad o disposición para resistirse a responder con la primera idea que nos viene a la mente” y propuso el Test de Reflexibilidad Cognitiva (TRC; Kahneman y Frederick, 2002) como un instrumento de medida de la RC. Desde entonces, autores de muy diversas disciplinas han mostrado un especial interés por examinar este constructo. Así, se ha sugerido que la RC está asociada con numerosos aspectos de nuestra vida diaria y que predice importantes resultados del día a día. Por ejemplo, se ha indicado que los sujetos con mayor RC obtienen mejores resultados en las tareas de toma de decisión, mostrando menor aversión al riesgo y mayor paciencia en el retorno de las recompensas (Campitelli y Labollita, 2010; Cokely y Kelley, 2009; Frederick, 2005). Además, comenten, en general, menos heurísticos y sesgos cognitivos en sus razonamientos (Hoppe y Kusterer, 2011; Kahneman, 2011; Moritz, Siemsen y Kremer, 2014; Sirota y Juanchich, 2011) y logran mejores comportamientos estratégicos (Brañas-Garza, García-Munoz y González, 2012). Asimismo, se ha evidenciado que son capaces de interpretar mejor las escenas de humor (Ventis, 2015), muestran menor tendencia a las creencias religiosas y a aspectos paranormales (Cheyne y Pennycook, 2013; Pennycook, Cheyne, Seli, Koehler y Fugelsang, 2012; Shenhav, Rand y Greene, 2011) y tienden a seguir menos las convicciones o valores morales (Aktaş, Yılmaz y Bahçekapılı, 2017; Baron, Scott, Fincher y Metz, 2015; Royzman, Landy y Leeman, 2015). Todos estos resultados han puesto de manifiesto la relevancia de la RC y la importancia de estudiar este constructo.

Aunque la investigación existente sobre la RC y el TRC es relativamente reciente, se han explorado diversas cuestiones al respecto. Así, la literatura sobre el tema se podría dividir en 3 grandes bloques. Por un lado, se encuentran los estudios que pretenden testar la validez de constructo del TRC. Es decir, tratan de responder a la pregunta ¿qué significa realmente la RC? a través de las preguntas ¿cuáles son las características del TRC?, ¿qué es lo que mide realmente el TRC?, ¿cómo lo hace? y ¿la forma en que lo hace es la correcta? Estos estudios se centran especialmente en testar tres grandes hipótesis o teorías: (1) El TRC es una medida más de capacidad cognitiva (Sinayev y Peters, 2015; Weller et al., 2013); (2) El TRC no solo es una medida de capacidad cognitiva sino que, además, mide capacidad o disposición hacia un

pensamiento racional (Federick, 2005; Toplak, West y Stanovich, 2011, 2014); y (3) el TRC mide capacidad cognitiva y la disposición de pensamiento minucioso, reflexivo y de cognición elaborada (Campitelli y Labollita, 2010; Cokely y Kelley, 2009).

Por otro lado, se encuentran los estudios que tienen por objetivo establecer los determinantes de la RC. La investigación se ha centrado en: (1) buscar qué variables individuales predicen el desempeño en el TRC y las diferencias individuales existentes en esas variables disposicionales y (2) establecer los determinantes situacionales que pueden tener un efecto directo en el desempeño en el TRC (Brañas-Garza, Kujal y Lenkei, 2015; Costa, Foucart, Arnon, Aparici y Apesteguia, 2014; De Neys, Rossi y Houdé, 2013; Mastrogiorgio y Petracca, 2014; Meyer et al., 2015; Simonovic, Stuppel, Gale y Sheffield, 2017). En este punto, también se encuentran las investigaciones que tienen por objetivo el “desesgo” (*debiasing*) del TRC para mejorar el rendimiento en el test.

La tercera línea de investigación se centra en examinar el procesamiento cognitivo que subyace en el proceso de respuesta al TRC. Aquí, se enmarcan los artículos sobre la meta-cognición de respuesta al test. Estos estudios se centran en explorar la consciencia de los sujetos a la hora de responder a los ítems del TRC (De Neys et al., 2013; Szaszi, Szollosi, Palfi y Aczel, 2017; Szollosi, Bago, Szaszi y Aczel, 2017; Trouche, Sander y Mercier, 2014).

A pesar de toda la literatura generada en torno a la RC, diversas cuestiones quedan todavía sin explorar o necesitan de más investigación para ser dilucidadas. Así, por ejemplo, es necesaria nueva investigación sobre la validez predictiva de la RC en el ámbito de la Psicología del Trabajo. A pesar de que se ha demostrado que la RC explica diversos resultados de nuestra vida cotidiana, apenas existe investigación que examine el impacto de la RC en el desempeño académico y ocupacional. Existen relativamente pocos estudios que exploran esta cuestión en el ámbito académico y, según nuestro conocimiento, solamente un estudio lo hace en relación con el desempeño ocupacional. No obstante, este tema toma especial relevancia en contextos ocupacionales donde el control de los impulsos y la autorregulación interna resultan de vital importancia (por ejemplo, una situación estresante o de emergencia). Del mismo modo, los estudios sobre la validez de constructo han arrojado resultados muy dispares y actualmente no existe un consenso sobre qué mide realmente el TRC. Nuevas investigaciones deben ser desarrolladas con el objetivo de aclarar este aspecto.

Algunos investigadores han planteado la necesidad de desarrollar nuevas medidas de RC (Salgado, Otero y Moscoso, 2019). El TRC (Frederick, 2005; Kahneman y Frederick, 2002) es el instrumento de evaluación de la RC por excelencia y, aunque existen nuevas versiones de este instrumento, el test de Frederick sigue siendo el más utilizado. El reducido número de ítems, el corto tiempo necesario para responder a los mismos y la aparente validez de la medida han provocado el incremento de su popularidad. No obstante, son diversas las limitaciones psicométricas que afectan a este test y que ponen en tela de juicio su validez como instrumento

de evaluación del constructo. En primer lugar, el reducido número de ítems y la dificultad de los mismos provoca una distribución sesgada de las puntuaciones y la incapacidad para discriminar entre los sujetos en función del nivel de rasgo latente. Estas características dificultan la aplicación de análisis bivariados y multivariados donde se asume la distribución normal de las puntuaciones. Asimismo, hacen problemática la estimación de la fiabilidad por consistencia interna. De este modo, el coeficiente de fiabilidad del instrumento tiende a ser bajo, con magnitudes inferiores a .70. Por otro lado, la gran popularidad del instrumento provoca una pérdida de validez de la medida, debido a la gran exposición de su contenido a la población. En conjunto, estas limitaciones han motivado la necesidad de crear nuevos tests de RC más válidos desde el punto de vista psicométrico.

Considerando lo anterior, el principal objetivo de esta tesis doctoral ha sido examinar la validez de constructo y de criterio de la medida de RC en el ámbito de la Psicología del Trabajo. Para ello, se ha tratado de validar un nuevo instrumento de RC que supere las limitaciones psicométricas anteriormente enumeradas. Los objetivos específicos que se plantearon fueron los siguientes: (1) evaluar las propiedades psicométricas de una nueva medida de RC; (2) conocer el efecto de diferentes aspectos relacionados con la aplicación de la prueba, a saber, la secuencia de presentación de los ítems o el formato de administración del test (lápiz y papel vs. ordenador); (3) mostrar evidencias de su validez basadas en los procesos de respuesta al test, analizando el tiempo de reacción a los ítems; (4) examinar la validez de constructo, estudiando la relación de la RC con medidas de capacidad mental general y variables de personalidad (*Big Five Factors* y disposiciones de pensamiento); (5) estimar la validez de criterio en relación con criterios académicos y ocupacionales; (6) analizar la validez añadida de la RC para predecir los criterios evaluados; (7) explorar las diferencias entre hombres y mujeres en RC; (8) estudiar la validez de constructo y de criterio de la RC en función del sexo; y (9) construir modelos explicativos de las relaciones halladas en los análisis anteriores.

Para conseguir los objetivos enumerados, esta tesis se dividió convenientemente en tres bloques. El primero de ellos, recoge el marco conceptual sobre el que se desarrollarán, posteriormente, los estudios empíricos. Esta primera parte se ha dividido en 3 capítulos teóricos, que se describen a continuación:

1. El primer capítulo teórico presenta un resumen de los principales Modelos Duales de Procesamiento (MDP) propuestos por las diferentes disciplinas. Estos modelos constituyen la base del desarrollo teórico del concepto de RC y del TRC. A lo largo del capítulo se examinan las características de los dos procesos cognitivos en los que se sustentan los modelos, se resumen sus principales funciones y la relación funcional entre ambos y se sintetizan las variables que modulan la activación de uno u otro proceso. Asimismo, se presentan brevemente otros modelos cognitivos que defienden la existencia

de un único proceso cognitivo o de múltiples formas de procesamiento o sistemas. Finalmente, se mencionan las recientes críticas efectuadas a la formulación teórica de los MPD.

2. El segundo capítulo se centra en introducir el concepto y el test de RC. Para ello, se define el constructo y se presentan los ítems que conforman el test, así como las características de corrección y administración de la medida. Por otro lado, se resume la literatura científica sobre el proceso cognitivo subyacente en la respuesta a los ítems y el proceso metacognitivo del que son conscientes los sujetos cuando responden correcta e incorrectamente a los ítems. Del mismo modo, también se sintetizan otros factores que pueden determinar la activación del procesamiento de tipo 1 o de tipo 2 en el desempeño en el TRC. Por último, en este capítulo se enumeran las principales limitaciones del TRC.
3. El tercer capítulo expone las principales evidencias de validez de constructo y de criterio de la medida de RC. El capítulo presenta, en primer lugar, una síntesis de las tres principales propuestas sobre qué mide el TRC: (a) capacidad cognitiva; (b) capacidad cognitiva y disposición hacia un pensamiento racional; y (c) capacidad cognitiva y disposición hacia un pensamiento de mente abierta. Seguidamente, se examina la relación de la RC con el modelo de personalidad de los *Big Five*, la variable sexo y la variable edad. Finalmente, se resumen las evidencias de validez de criterio. Específicamente, se examinan los limitados estudios existentes sobre la relación de la RC con resultados académicos y ocupacionales. Asimismo, se detallan las evidencias de validez añadida de la RC para predecir dichos criterios.

Una vez revisada la parte teórica, la segunda parte de este trabajo presenta los estudios empíricos llevados a cabo para explorar cada uno de los objetivos anteriormente propuestos y las hipótesis planteadas al respecto. En este sentido, la tesis se compone de 4 estudios empíricos que se describen seguidamente:

1. El objetivo del primer estudio ha sido evaluar la estructura interna de un nuevo test de RC (Salgado, 2014a) y comparar los resultados con los hallados para el TRC original (Frederick, 2005; Kahneman y Frederick, 2002). Para ello, se han administrado ambos tests a una muestra de 1,367 sujetos. Se estiman los estadísticos descriptivos y la estructura factorial del nuevo test y la fiabilidad por consistencia interna, test-retest y equivalencia y estabilidad de ambos instrumentos de evaluación. Los resultados ponen en evidencia las limitaciones del TRC original y las dificultades que estas implican a la hora de examinar sus propiedades psicométricas, como, por ejemplo, el análisis de la fiabilidad. Del mismo modo, los resultados muestran que el nuevo TRC es un instrumento adecuado de evaluación de la RC y que es superior, psicométricamente, al original. Las puntuaciones del test se distribuyen normalmente y, además, presentan un nivel adecuado de fiabilidad. Finalmente, este estudio también muestra la existencia de diferencias en

- función del sexo en el TRC, siendo los hombres los que alcanzan mejores resultados en el test.
2. El segundo estudio se ha llevado a cabo con el objetivo de examinar el efecto que diferentes factores relacionados con la aplicación y administración del test pueden tener sobre el desempeño en el nuevo TRC. Concretamente, se exploró si manipular el orden de presentación de los ítems afecta a los resultados y al tiempo invertido en responder al nuevo TRC en general, y a los ítems originales o a los nuevos, en particular. En segundo lugar, se examinó la relación entre el tiempo de respuesta y la puntuación alcanzada en el TRC. Finalmente, como los resultados del estudio 1 han reflejado diferencias en función del sexo, también se han examinado los efectos de forma independiente en hombres y mujeres. Los resultados han sugerido que manipular la secuencia de presentación de los ítems afecta al desempeño en el TRC y, únicamente, al tiempo invertido en responder a los primeros ítems del test. No obstante, los nuevos ítems controlan mejor los efectos que produce el orden de presentación de los mismos en el desempeño en el test. Asimismo, se ha encontrado una asociación positiva entre la RC y el tiempo invertido en responder.
 3. El propósito del tercer estudio ha sido doble. (1) Por un lado, se ha pretendido conocer la validez de constructo de la medida de RC. Para ello, se ha examinado la validez convergente y discriminante utilizando variables cognitivas (capacidad mental general) y de personalidad (*Big Five Factors* y disposiciones de pensamiento). Además, se ha examinado la validez conjunta de todas esas variables individuales para determinar la RC. Como resultado, se han propuesto diversos modelos explicativos del desempeño en el TRC. (2) El segundo objetivo ha sido estimar la validez predictiva de la RC. Se han examinado criterios académicos y ocupacionales. Del mismo modo, se ha investigado la validez que la RC añade sobre otras variables individuales en la predicción del desempeño. Finalmente, las evidencias de validez se han mostrado por separado para el grupo de hombres y mujeres. Los resultados han puesto de manifiesto que la RC es un constructo principalmente cognitivo, aunque también se relaciona con variables de personalidad. Asimismo, se ha observado que, aunque las variables disposicionales no determinan del mismo modo la RC en hombres y mujeres, en su conjunto explican el mismo porcentaje de varianza en ambos sexos. Por último, los datos evidencian la validez de la RC para predecir el desempeño académico y ocupacional con independencia de la variable sexo. Además, se ha encontrado que la RC añade validez sobre otras variables individuales en la predicción del desempeño ocupacional y académico.
 4. El cuarto estudio presenta tres meta-análisis de la RC. El primero de ellos, examina la relación entre la capacidad cognitiva y la RC y comprueba si esta relación se generaliza a la población. El segundo examina la validez teórica y operativa de la RC para predecir el desempeño académico y si se generaliza su validez a la población. Finalmente, el tercer

meta-análisis examina las diferencias en función del sexo. Las tres integraciones cuantitativas se han llevado a cabo nuevamente examinando el tipo de TRC (TRC original y nuevas versiones del test) como posible variable moderadora. Los resultados apoyan los hallazgos encontrados en los tres estudios previos sobre la relación entre la RC y las tres variables examinadas.

Por último, el tercer bloque en que se divide este trabajo presenta las conclusiones generales de la tesis doctoral y las contribuciones que cada estudio empírico ha efectuado al conocimiento de la RC. Del mismo modo, se enumeran las implicaciones prácticas que los resultados tienen en el dominio de la Psicología del Trabajo y se describen las principales líneas de investigación para futuros estudios.



MARCO **C**ONCEPTUAL





CAPÍTULO 1.

Modelos de Procesamiento Dual (*Dual-Process Models*)

1. INTRODUCCIÓN

Tradicionalmente, se ha afirmado que el ser humano se comporta de forma racional o de acuerdo con lo que es más beneficioso para sí mismo (Strack y Deutsch, 2004). Sin embargo, estudios sobre el razonamiento han mostrado que el resultado observado es frecuentemente deficiente, en el sentido de que en el procesamiento cognitivo se cometen errores, y altamente inconsistente a lo largo de las situaciones, al mostrar un amplio abanico de conductas en función del sujeto y la situación planteada (Evans, 1984). Consecuentemente, estos resultados han creado la necesidad de explorar y explicar las posibles causas del comportamiento.

Varios estudios han sugerido que el razonamiento humano podría estar guiado por dos procesos cognitivos diferentes que interactúan entre sí (Chaiken y Trope, 1999; Evans y Over, 1996; Frankish, 2010; Kahneman y Frederick, 2002, 2005; Stanovich, 1999, 2011). Estas investigaciones han desencadenado la formulación de modelos de procesamiento dual (en adelante MPD) en múltiples disciplinas. Dentro de la psicología cognitiva se han desarrollado modelos aplicados al contexto del razonamiento humano, los mecanismos generales de pensamiento y los sesgos cognitivos (Evans, 1984, 1989, 2006; Evans y Wason, 1976; Pollock, 1989, 1991; Reber, 1993; Schneider y Shiffrin, 1977). En la psicología social se han desarrollado MPD para explicar, por ejemplo, el cambio de actitudes, los estereotipos, la creación de los juicios sociales, el control de los impulsos y la autorregulación (Bazerman, Tenbrunsel y Wade-Benzoni, 1998; Chaiken, Liberman y Eagly, 1989; Hammond, 1996; Loewenstein, 1996; Metcalfe y Mischel, 1999). En economía, los modelos han sido aplicados al dominio del consumo y la toma de decisiones (Bazerman et al., 1998; Hammond, 1996; Kahneman y Frederick, 2002, 2005; Thaler y Shefrin, 1981).

Los MPD han sido desarrollados por sus autores de forma independiente, frecuentemente, ignorando la existencia de las publicaciones de otros investigadores (Evans y Stanovich, 2013).

Sin embargo, aunque con sutiles diferencias entre ellos, todos defienden que el funcionamiento de nuestro cerebro está caracterizado por una dicotomía de procesos cognitivos. Proponen que los dos tipos de procesamiento pueden ser definidos uno como automático e intuitivo, que opera de forma rápida y con poco o ningún esfuerzo, y otro más lento, deliberado y reflexivo, que opera con esfuerzo y con concentración (Frankish, 2010; Guthrie, Rachlinski y Wistrich, 2007; Haidt, 2001, 2008; Spunt, 2015; Stanovich, 2011). Ejemplos de procesamientos intuitivos o automáticos serían completar la frase “ $2 + 2 = ?$ ”, detectar enfado en la voz de una persona o percibir si un objeto está más lejos que el otro (Kahneman, 2011). Sin embargo, ejemplos de procesamientos reflexivos serían estar atentos al disparo de salida de una carrera, hacer operaciones matemáticas complejas o contar el número de letras *a* que hay, exactamente, en esta página (Kahneman, 2011).

Aunque existe un amplio consenso en cuanto a diferenciar dos tipos de procesamiento y las características, funciones y procesos generales que implican, este acuerdo sigue suscitando dudas entre algunos autores (Glöckner y Witteman, 2010). Así, varios investigadores han propuesto modelos cognitivos que postulan un procesamiento único (Kruglanski y Gigerenzer, 2011) o más de dos procesos o sistemas cognitivos diferentes (Conrey, Sherman, Gawronski, Hugenberg y Groom, 2005; Evans, 2008; Glöckner y Witteman, 2010; Stanovich, 2011). Incluso se han propuesto modelos que defienden un continuo de los procesos, donde el procesamiento intuitivo y analítico conforman los polos opuestos de ese continuo (Hammond, 1996; Reber, 1993). A pesar de estas excepciones, la mayor parte de las teorías comparten las propiedades generales de los dos tipos de procesamiento.

El objetivo de este primer capítulo es analizar la literatura sobre los dos tipos de procesamiento cognitivo, tratando de unificar los MPD desarrollados en las diferentes disciplinas. Para ello, en primer lugar, se presentan las características de ambos procesos, sus principales funciones, la relación funcional entre ambos y las variables que modulan la activación de uno u otro. A continuación, se describen brevemente otros modelos cognitivos que defiende la existencia de un único proceso cognitivo o múltiples formas de procesamiento o sistemas. Por último, se apuntan las recientes críticas efectuadas a la formulación teórica de los MPD.

2. MODELOS DE PROCESAMIENTO DUAL (MPD)

Los MPD se fundamentan en la idea de que el razonamiento humano está controlado por dos tipos de procesamientos cognitivos cualitativamente diferentes (Strack y Deutsch, 2004). Tradicionalmente, se ha identificado como sistema 1 (*system 1*) y sistema 2 (*system 2*) al procesamiento intuitivo y analítico, respectivamente. Esta terminología, propuesta por Stanovich (1999), ha sido la más ampliamente utilizada y aceptada por ser considerada más neutra y no hacer referencia a ninguna característica o tributo propio de cada tipo de procesamiento (Evans,

2008). Recientemente, Evans y Stanovich (2013; Evans, 2008; Stanovich, 2011) han afirmado que esta denominación podría suscitar problemas. En primer lugar, hablar de sistemas duales puede ser ambiguo, ya que, en ocasiones, esa nomenclatura es empleada para referirse a la hipótesis de la existencia de dos mentes que va más allá de la diferenciación entre dos tipos de procesamiento (por ejemplo, Bechara, 2005; Kahneman, 2011; Stanovich, 2011). La hipótesis de la existencia de dos mentes defiende la convivencia de dos sistemas, con marcadas diferencias evolutivas y estructuras neurológicas diferentes, comúnmente denominados sistema 1 y sistema 2, que sustentan los dos tipos de procesamiento, procesamiento de tipo 1 y procesamiento de tipo 2, respectivamente (Evans, 2006, 2008; Frankish, 2010). Esta hipótesis ha dado lugar a lo que se conoce como Modelo Dual de Sistemas (*Dual-System Model*). Así, es frecuente encontrar modelos que hacen referencia a un Modelo Dual de Sistemas (en adelante MDS) para describir, realmente, a un MPD (Evans y Stanovich, 2013; Evans, 2008). Por ejemplo, el modelo propuesto por Sloman (1996) denominó sistema asociativo al procesamiento de tipo 1 y sistema basado en reglas al procesamiento de tipo 2. Epstein (2003) identifica al procesamiento de tipo 1 como sistema racional y al procesamiento de tipo 2 como sistema experiencial. Strack y Deutsch (2004) proponen un sistema impulsivo y un sistema reflexivo para referirse a un procesamiento de tipo 1 y de tipo 2, respectivamente. Por último, Guthrie y colaboradores (2007) diferencian entre sistema intuitivo, para hacer referencia al procesamiento de tipo 1 y sistema deliberado para referirse al procesamiento de tipo 2.

En segundo lugar, los términos sistema 1 y sistema 2 están formulados en singular. Sin embargo, recientemente, se ha sugerido que ambos sistemas podrían estar compuestos por un conjunto de sistemas (Evans, 2008; Stanovich, 2011) o que ambos tipos de procesamiento cognitivos podrían englobar múltiples procesos en su funcionamiento (Modelos de Procesos Múltiples; Conrey et al., 2005; Glöckner y Witteman, 2010).

Consecuentemente, Evans y Stanovich (2013) propusieron retomar la antigua nomenclatura de Evans y Wason (1976; Wason y Evans, 1975), recomendando el uso de procesamiento de tipo 1 (*type 1 processing*) para referirse a un procesamiento intuitivo, y procesamiento de tipo 2 (*type 2 processing*) para referirse a un procesamiento analítico. Los autores admiten que esta terminología es menos ambigua y no excluye la posible connotación de que el procesamiento cognitivo podría resultar de la operación de una variedad de sistemas y procesos (Evans y Stanovich, 2013; Evans, 2008; Stanovich, 2011). Por ello, a lo largo de esta tesis, emplearemos las denominaciones procesamiento de tipo 1 y procesamiento de tipo 2 para referirnos a un procesamiento intuitivo y analítico, respectivamente.

En cualquier caso, se han propuesto otras nomenclaturas durante el desarrollo de los diversos MPD. En la Tabla 1 se recoge una lista no exhaustiva de los principales modelos de

procesamiento existentes (MPD y MDS). Las dos primeras columnas recogen las diversas terminologías propuestas para denominar a cada procesamiento o sistema cognitivo. Por un lado, se muestran las calificaciones acuñadas originalmente en inglés y, por otro, su correspondiente traducción al español. Asimismo, la tabla sintetiza brevemente cada modelo, haciendo mención al ámbito de aplicación y a sus características más relevantes.

2.1. Sistemas Asociados a los Procesamientos

Como se ha señalado anteriormente, la hipótesis de dos mentes propone que los dos tipos de procesamiento están asociados a sistemas cognitivos separados (Evans y Over, 1996; Kahneman y Frederick, 2002, 2005; Lieberman, 2003; Metcalfe y Mischel, 1999; Stanovich, 1999, 2011). Esta hipótesis sugiere que el sistema vinculado al procesamiento de tipo 2 es únicamente humano y que es la fuente de capacidad para llevar a cabo pensamientos hipotéticos y abstractos. En su forma más ambiciosa, los MDS sugieren que los humanos disponemos de dos mentes, con estructuras neurológicas diferentes (Frankish, 2010).

Se han propuesto diversos modelos donde cada tipo de procesamiento es asociado a un sistema concreto. Así, Evans y Over (1996) presentaron dos tipos de racionalidad: una personal (racionalidad₁) y otra impersonal (racionalidad₂) que asociaron a dos sistemas de procesamiento diferentes: sistema implícito y explícito, respectivamente. La racionalidad personal hace referencia a cuáles son nuestros objetivos individuales y si estamos actuando para tratar de alcanzarlos. El razonamiento impersonal hace referencia a si estamos siguiendo los principios de la lógica y otras teorías normativas en nuestro razonamiento y toma de decisiones. El sistema implícito, también denominado tácito o inferencial, está unido con la racionalidad₁ y está altamente enfocado hacia las características relevantes de la información de la tarea y de la memoria. Además, tiende a aplicar el conocimiento general a casos específicos. Por su parte, la racionalidad₂ está unida al sistema explícito. Este sistema está asociado con el razonamiento verbal e influido por las instrucciones verbales. Además, está altamente enfocado hacia la información concreta y sus funciones complementan las limitaciones del sistema implícito. El sistema explícito se encarga de gestionar los casos novedosos y se anticipa al futuro.

Tabla 1.

Lista no Exhaustiva de los Principales Modelos de Procesamiento y Nomenclatura de los Procesamientos o Sistemas Propuestos

Autores		Procesamiento de Tipo 1	Procesamiento de Tipo 2	Características relevantes del Modelo de Procesamiento
1.	Wason y Evans (1975; Evans y Wason, 1976)	Proceso de tipo 1 <i>Type 1 processes</i>	Proceso de tipo 2 <i>Type 2 processes</i>	Centrado en el ámbito cognitivo del razonamiento inductivo y deductivo. Desarrollo del modelo en base a los resultados obtenidos en estudios sobre la tarea de selección de Wason (1966).
2.	Schneider y Shiffrin (1977; Shiffrin y Schneider, 1977)	Procesamiento automático o detección automática <i>Automatic processing or automatic detection</i>	Procesamiento controlado o búsqueda controlada <i>Controlled processing or controlled search</i>	Centrado en el ámbito cognitivo del funcionamiento de la mente. Desarrollo del modelo en base al funcionamiento de la memoria a corto y largo plazo formados por nodos activados automática o temporalmente según el proceso.
3.	Chaiken, Liberman y Eagly (1989; Chaiken, 1980)	Procesamiento heurístico <i>Heuristic processing</i>	Procesamiento sistemático <i>Systematic processing</i>	Aplicado al ámbito de la psicología social para explicar la persuasión en la comunicación e interpretación o evaluación de los mensajes. Desarrollo del modelo heurístico-sistemático de procesamiento, donde ambos tipos de procesamiento pueden concurrir u operar independientemente para juzgar la validez del mensaje y, los heurísticos o sesgos cognitivos pueden operar deliberada o automáticamente.
4.	Thaler y Shefrin (1981)	Enprendedor miope <i>Myopic doer</i>	Planificador predictor <i>Farsighted planner</i>	Aplicado al contexto económico del consumo. Proponen un modelo dicotomizado de procesamiento para explicar el autocontrol de los sujetos ante el conflicto entre las preferencias a corto y largo plazo: promover un beneficio inmediato o un beneficio mayor a largo plazo.
5.	Evans (1984, 1989, 2006)	Procesamiento heurístico (o implícito) <i>Heuristic (or implicit) processing</i>	Procesamiento analítico (o explícito) <i>Analytic (or explicit) processing</i>	Centrado en el ámbito cognitivo del razonamiento inductivo y deductivo. Evans (1984, 1989) propone la teoría de razonamiento heurístico-analítico basada en la idea de que la irracionalidad deriva de la comisión de sesgos. Posteriormente, añade tres principios extraídos de la teoría de pensamiento hipotético de Evans, Over y Handley (2003), donde introduce la nomenclatura de procesamiento implícito/explicito (Evans, 2006).

Continúa

Tabla 1.
Continuación

Autores		Procesamiento de Tipo 1	Procesamiento de Tipo 2	Características relevantes del Modelo de Procesamiento
6.	Norman y Shallice (1986)	Planificación de los argumentos <i>Contention scheduling</i>	Sistema atencional de supervisión <i>Supervisory attentional system</i>	Proponen un modelo teórico conducente a explicar el rol de la atención en la selección y control de las acciones o conductas de los sujetos, cuando el desempeño es automático o deliberado.
7.	Logan (1988)	Procesamiento automático <i>Automatic processing</i>	Procesamiento no automático o algorítmico <i>Nonautomatic or algorithmic processing</i>	Presenta la teoría de los casos de automatización (<i>Instances Theory of Automatization</i>). Esta teoría asume que el procesamiento automático está construido sobre la adquisición de conocimiento procedente del entorno y de nuestra consistente experiencia con el mismo. La información se almacena en forma de representaciones (denominadas <i>instances</i>) que serán recuperados para aplicar en ocasiones similares (consistencia de la experiencia).
8.	Pollock (1989; 1991)	Procesos no intelectuales (razonamiento rápido e inflexible) <i>Nonintellectual processes (Quick and inflexible modules)</i>	Intelección <i>Intellection</i>	Desarrolla la teoría general del razonamiento (OSCAR) que diferencia entre procesos no intelectuales y el proceso de la intelección, que a su vez se divide en dos tipos de racionalidad (teórica y práctica). Este modelo no sugiere un procesamiento dicotómico, sino que subdivide el proceso deliberado.
9.	Reber (1993)	Proceso implícito o conocimiento tácito <i>Implicit process and tacit knowledge</i>	Proceso explícito <i>Explicit process</i>	Aplicado al dominio del razonamiento humano. Propone un modelo continuo (o no dicotómico) del procesamiento humano, al asumir que no existe una clara delimitación entre el procesamiento implícito y explícito. Ambos tipos de cognición se encuentran en los polos del continuo y el resto de los procesos incluyen elementos del pensamiento consciente e intuitivo.

Continúa

Tabla 1.
Continuación

Procesamiento		Características relevantes del Modelo de Procesamiento	
de Tipo 1		de Tipo 2	
Autores			
10. Loewenstein (1996)	Control visceral (Factores viscerales)	Control cognitivo (Intereses)	Aplicado al contexto social de autorregulación y control de los impulsos en conductas autodestructivas (consumo drogas, consumo abusivo de alimentos, malos hábitos sexuales, comisión de crímenes). El modelo trata de explicar las discrepancias entre los intereses de los sujetos (buena salud) y sus comportamientos finales (consumo drogas). Trata de explicar la regulación de las acciones a través de los factores viscerales.
	<i>Visceral control (Visceral factors)</i>	<i>Cognitive control (Interest)</i>	
11. Sloman (1996)	Sistema asociativo	Sistema basado en reglas	Centrado en el ámbito cognitivo del razonamiento humano. Presenta un modelo dual de sistemas de razonamiento y solución de problemas. Sugiere la posibilidad de que ambos sistemas propongan soluciones a un problema que sea contradictorias entre sí, y que el sujeto crea simultáneamente en ambas posibilidades (criterio S).
	<i>Associative system</i>	<i>Rule-based system</i>	
12. Hammond (1996)	Cognición intuitiva	Cognición analítica	Centrado en la psicología social en el dominio de la toma de decisión y juicio. Propone un modelo no dicotómico, es decir, sugiere que el procesamiento es un continuo que se mueve entre el polo analítico e intuitivo, donde diversas combinaciones dan lugar a múltiples procesos (quasi-racionalidad). El proceso de “sentido común” es el más utilizado habitualmente.
	<i>Intuitive cognition</i>	<i>Analytic cognition</i>	
13. Evans y Over (1996)	Proceso de pensamiento (o sistema) tácito o implícito (racionalidad ₁)	Proceso de pensamiento (o sistema) explícito (racionalidad ₂)	Enfocado al ámbito del razonamiento humano. El modelo diferencia entre dos tipos de racionalidad, y asocia cada una de ella a un sistema/proceso de pensamiento. Sugieren una distinción entre la teoría dual de procesamiento y la teoría dual de sistemas. Este modelo contempla la comisión de sesgos cognitivos en el procesamiento cognitivo.
	<i>Tacit or implicit thought processes or system (rationality₁)</i>	<i>Explicit thought processes or system (rationality₂)</i>	
14. Bazerman, Tenbrunsel y Wade-Benzoni (1998)	Yo quiero	Yo debería	Aplicado a la psicología social, al ámbito de la toma de decisión en la vida diaria. Proponen un modelo dicotómico de procesamiento para explicar el conflicto intrapersonal que surge entre lo que yo quiero/desco hacer y lo que debería hacer, cuando tomo una decisión.
	<i>Want self</i>	<i>Should self</i>	

Continúa

Tabla 1.
Continuación

Procesamiento		Características relevantes del Modelo de Procesamiento	
Autores	de Tipo 1	Procesamiento de Tipo 2	
15. Metcalfe y Mischel (1999)	Sistema “cálido” o Sistema del conocimiento <i>Hot system or Know system</i>	Sistema “frío” o Sistema de ir <i>Cool system or Go system</i>	Centrado en la psicología social, en el ámbito del autocontrol, fuerza de voluntad o autorregulamiento en el aplazamiento o demora de las gratificaciones. Se trata de un modelo dicotómico de sistemas compuestos por nodos que interactúan entre ellos. Sugieren la posibilidad de que el sistema “cálido” domine el funcionamiento del sistema “frío”.
16. Bargh y Chartrand (1999)	Proceso mental automático o no consciente <i>Automatic or non-conscious mental process</i>	Proceso mental consciente o deliberado <i>Conscious or deliberate mental process</i>	Aplicado al dominio del razonamiento humano. Proponen un modelo dual de procesamiento. Definen las características de ambos procesos. Sugieren que el proceso automático puede funcionar de forma natural y no solo como consecuencia de la consistencia de la experiencia (aprendizaje). Afirman que muchas tareas cotidianas son llevadas a cabo mediante el proceso no consciente.
17. Stanovich (1999, 2011)	Sistema 1 <i>System 1 (TASS)</i>	Sistema 2 <i>System 2</i>	Enfocado al dominio del razonamiento humano. Propone un modelo dual de sistemas de procesamiento (Stanovich, 1999), posteriormente (Stanovich, 2011) propone una extensión tripartita del modelo y asume que el sistema 1 no está compuesto por un único sistema sino por un conjunto de ellos denominado TASS y que el sistema 2 se puede dividir en una mente reflexiva y una mente algorítmica. Finalmente, asocia cada sistema de procesamiento con un tipo de inteligencia concreta (Stanovich, 1999, 2011; Stanovich y West, 2000)
18. Smith y DeCoster (2000)	Procesamiento asociativo <i>Associative processing mode</i>	Procesamiento basado en reglas <i>Rule-based processing mode</i>	Proponen un modelo dual de procesamiento y un modelo dual de sistemas de memoria aplicados al ámbito de la psicología social y cognitiva. Además, hacen una revisión de otras teorías duales en el mismo ámbito. Asumen que diversos factores pueden moderar la activación de los sistemas (motivación y carga cognitiva). Integran, por tanto, nuevas variables en el modelo dual que permiten explicar la <i>akrasia</i> , es decir, la falta de auto control o la tendencia a actuar en contra del mejor juicio.

Continúa

Tabla 1.
Continuación

Procesamiento de Tipo 1		Procesamiento de Tipo 2	Características relevantes del Modelo de Procesamiento
19. Brainerd y Reyna (2001)	Procesamiento del quiz de la cuestión (memorias del quiz de la cuestión)	Procesamiento difuso (memoria verbatim)	Aplicado al área del razonamiento humano. Proponen la teoría del Trazo Difuso (<i>Fuzzy-Trace Theory</i>) para explicar el procesamiento dual de la memoria y el razonamiento. La teoría postula que, aunque la memoria y el razonamiento se encuentran ligados, hay situaciones en las que operan de manera disociada. No obstante, esta teoría, a diferencia de los modelos duales clásicos, se sustenta sobre la base de que se debe promover el razonamiento intuitivo sobre el razonamiento analítico. Este punto es clave para comprender las propuestas del modelo.
	<i>Gist processing (Gist memories)</i>	<i>Fuzzy processing (verbatim memories)</i>	
20. Haidt (2001, 2008)	Sistema intuitivo	Sistema racional	Aplicado al dominio de la creación de juicios morales dentro de la psicología social. Propone un modelo social intuitivo para explicar los juicios morales. Se construye en base a la idea de que los juicios morales son causados por intuiciones morales que se generan rápida y automáticamente. Por lo tanto, aplica las características de los dos tipos de procesamiento para explicar la configuración de los juicios morales.
	<i>Intuitive system</i>	<i>Reasoning system</i>	
21. Nisbett, Peng, Choi y Norenzayan (2001)	Pensamiento holístico	Pensamiento analítico	Aplicado al ámbito de la psicología social. Los autores aplican las características de los MPD para explicar las diferencias culturales entre sujetos de nacionalidad griega y de nacionalidad china.
	<i>Holistic thought</i>	<i>Analytic thought</i>	
22. Kahneman y Frederick (2002, 2005)	Sistema 1	Sistema 2	Aplicado al dominio de la configuración de juicios, solución de problemas o toma de decisiones dentro de la psicología social. Proponen un modelo dual de sistemas de procesamiento, adoptando la inicial nomenclatura de Stanovich (1999). Proponen nuevos moduladores de la activación de ambos sistemas. Afirmar que el mayor factor explicativo de la formación de juicios son los sesgos o heurísticos (atajos cognitivos), que se aplican cuando ambos sistemas fallan en su funcionamiento.
	<i>Intuición</i>	Razonamiento	
	<i>System 1</i>	<i>System 2</i>	
	<i>Intuition</i>	<i>Reasoning</i>	

Continúa

Tabla 1.
Continuación

Autores		Procesamiento de Tipo 1	Procesamiento de Tipo 2	Características relevantes del Modelo de Procesamiento
23.	Epstein (2003)	Sistema experiencial <i>Experiential system</i>	Sistema racional <i>Rational system</i>	Aplicado al dominio del psicoanálisis o psicoterapia. Epstein propone la Auto-teoría Cognitiva-experiencial (<i>Cognitive-Experiential Self-Theory</i> ; CEST) definida como una teoría integrativa de la personalidad, que diferencia entre dos sistemas de procesamiento.
24.	Lieberman (2003)	Proceso reflejo <i>Reflective process</i>	Proceso reflexivo <i>Reflective process</i>	Enfocado al contexto de los juicios dentro de la psicología social y la neurociencia. Desarrollan un modelo dicotómico, reflejo-reflexión, de sistemas y tipos de procesamiento. Cada tipo de procesamiento está asociado a un sistema neuronal diferente, que activa distintas zonas cerebrales en su funcionamiento.
25.	Strack y Deutsch (2004)	Sistema impulsivo <i>Impulsive system</i>	Sistema reflexivo <i>Reflective system</i>	Modelo enfocado al comportamiento o conducta dentro de un contexto social. Proponen un modelo dual de sistemas que trata de explicar la <i>akrasia</i> dentro del contexto social. El modelo asume que el comportamiento social es el resultado del funcionamiento de ambos sistemas. Asimismo, sugiere que diversos elementos motivacionales conducen el comportamiento humano, que termina siendo inconsistente con los valores y creencias de la persona. Para explicar el vacío temporal que transcurre entre la toma de decisión y la implementación de las acciones, el modelo introduce mecanismos de intención.
26.	Bechara (2005)	Sistema impulsivo <i>Impulsive system</i>	Sistema reflexivo <i>Reflective system</i>	Perspectiva neurocognitiva aplicada dentro del contexto social de toma de decisión y control de impulsos en el consumo de drogas. Presenta un modelo dual de sistemas de procesamiento estableciendo las estructuras neuro-cerebrales que operan en cada sistema. Identifica a la amígdala como región clave del procesamiento impulsivo y al córtex prefrontal ventromedial con el proceso reflexivo.
27.	Gawronski y Bodenhausen (2006)	Proceso asociativo implícitas <i>Associative processes</i>	Proceso proposicional (actitudes explícitas) <i>Propositional processes</i>	Desarrollada dentro de la psicología social para explicar el cambio de actitudes en la evaluación de entidades. Los autores proponen un nuevo modelo dual de procesamiento, denominado modelo de evaluación asociativo-proposicional (<i>associative-propositional evaluation model</i>), para explicar los cambios de actitudes, implícitas y explícitas, en los juicios evaluativos de objetos. Asocia un tipo de actitud a cada proceso cognitivo en función de su creación.

Continúa

Tabla 1.
Continuación

Autores		Procesamiento de Tipo 1	Procesamiento de Tipo 2	Características relevantes del Modelo de Procesamiento
28.	Toates (2006)	Proceso basado en los estímulos <i>Stimulus-based process</i>	Proceso de orden superior <i>Higher order process</i>	Presenta una propuesta de modelo dual de procesamiento con un enfoque biológico, aunque aplicado al contexto del comportamiento social. Le da importancia al desarrollo evolutivo de los dos procesos. Asocia cada proceso cognitivo con conexiones neuronales diferentes e identifica las regiones del cerebro implicadas en cada uno de ellos. Asimismo, señala ciertos factores que funcionan de moduladores de la activación de los procesos de pensamiento.
29.	Guthrie, Rachlinski y Wistrich (2007)	Sistema intuitivo <i>Intuitive system</i>	Sistema deliberado <i>Deliberative system</i>	Proponen un modelo de juicio intuitivo-anulado (<i>intuitive-override judgment model</i>) basado en la teoría dual de procesamiento que trata de explicar la toma de decisión y los juicios adoptados por los jueces en un contexto judicial o legal. El modelo asume muchas semejanzas con los modelos duales propuestos por Kahneman y Frederick (2002, 2005) y Sloman (1996). Llevaron a cabo estudios que evidencian los dos tipos de procesamiento cognitivo a través de la aplicación del Test de Reflexividad Cognitiva (TRC; Frederick, 2005).

Metcalfe y Mischel (1999) propusieron un modelo dicotómico de sistemas de procesamiento que diferencia entre el sistema “frío” (*cool system*) y el sistema “cálido” (*hot system*). El primero está orientado a un procesamiento de tipo 2 y es emocionalmente neutro, flexible, coherente, lento y estratégico y es el responsable del autocontrol y la autorregulación de las respuestas. El sistema “cálido” está más orientado a un procesamiento de tipo 1 y es definido como simple y rápido. Es la base de las emociones, los miedos y las pasiones. Está especializado en el procesamiento automático de las emociones y responde en base a la activación de las representaciones previamente aprendidas y guardadas en la memoria. Metcalfe y Mischel (1999) afirman que ambos sistemas están organizados en nodos internos. Los nodos del sistema “cálido” se pueden identificar como fragmentos de sentimientos y no se encuentran conectados entre sí, de modo que su funcionamiento es cíclico.

Sin embargo, los nodos del sistema “frío” se encuentran interconectados entre sí y, a su vez, conectados con los del sistema “cálido”. Cuando opera el sistema “cálido”, se activan las relaciones afectivas y emocionales relacionadas con el estímulo en cuestión. Cuando opera el sistema “frío”, se activan los nodos que lo conforman, encargados de guardar toda la información relacionada con el contexto, las consecuencias y demás características de la situación que permiten su relación con otros conceptos y rasgos, proveyendo conocimiento acerca del estado y la situación. La interconexión entre los diversos nodos del sistema “frío” permiten el control de los impulsos del sistema “cálido”.

Lieberman (2003) también presentó un MPD. Este autor propuso una dicotomía de procesamiento de juicios que denominó proceso reflejo y proceso reflexivo para referirse a un procesamiento automático y controlado, respectivamente. El modelo de procesamiento reflejo-reflexivo establece que cada tipo de procesamiento se encuentra asociado con un sistema neural concreto. Así, el proceso reflejo se asocia con el sistema neural X y el proceso reflexivo con el sistema neural C. El sistema X es el responsable de asignar a cada estímulo su significado afectivo y social correspondiente. Esto incluye la asociación de estereotipos y categorías. El sistema C es el encargado de operar cuando el proceso reflejo falla y, por ende, se activa el proceso reflexivo.

Finalmente, Stanovich (1999, 2011) identificó los términos sistema 1 y sistema 2 para referirse al procesamiento de tipo 1 y de tipo 2, respectivamente. Posteriormente, Kahneman y Frederick (2002, 2005) siguieron esta misma terminología para referirse a un procesamiento intuitivo y deliberado, respectivamente. De modo que esta nomenclatura ha sido la más utilizada.

2.2. Memorias Asociadas a los Procesamientos

Algunos MPD han asociado los procesamientos cognitivos con un tipo de memoria específica. Schneider y Shiffrin (1977; Shiffrin y Schneider, 1977) describen la memoria como

un conjunto de nodos que se encuentran interrelacionados de manera compleja y permanente a través del aprendizaje. Algunos de estos nodos se encuentran inactivos y almacenados en una sección de la memoria a largo plazo, los otros se encuentran activos y almacenados en la memoria a corto plazo. La primera guarda información de forma permanente mientras que la segunda lo hace de forma temporal. Cuando tratamos de enviar información de la memoria a corto plazo a la memoria a largo plazo, esta se puede perder o simplemente olvidar. Cuando tratamos de manipular los flujos de información tanto dentro como fuera de la sección a corto plazo estamos aplicando un proceso consciente y deliberado (procesamiento de tipo 2). Asumimos el control del procesamiento de la información y esta información es guardada en la sección a largo plazo. Por lo tanto, todos los patrones y secuencias aprendidos se guardan en la sección a largo plazo que, previamente, han sido ejecutados automáticamente por la memoria a corto plazo. Los autores sugieren la existencia de dos procesos cognitivos de pensamiento: uno automático y otro controlado (Shiffrin y Schneider, 1977). Asumen que el proceso automático (*automatic process*) está formado por una secuencia de nodos que se activan automáticamente en respuesta a diversas entradas (*inputs*) generadas interna o externamente. Cuando la secuencia se inicia, sus nodos son activados y la información asociativa pasa de la memoria a largo plazo (aprendida) a la sección de la memoria a corto plazo. Sin embargo, los autores sugieren que el proceso controlado (*controlled process*) se inicia cuando se activa la secuencia de nodos temporales controlados bajo la atención del sujeto y solo una parte de ella conseguirá ser transferida permanentemente a la sección a largo plazo sin ser perdida u olvidada durante el proceso. De este modo, la sección a corto plazo mantiene dos roles diferenciados: (1) abastece al procesamiento de tipo 2 de la información temporal que es relevante para su procesamiento, es decir, reduce la cantidad de información que recibe en cantidades más manejables y (2) deja espacio para la toma de decisión, el razonamiento y cualquier otro proceso controlado por el sujeto. Por lo tanto, en ocasiones, el proceso controlado es iniciado por un proceso automático, especialmente cuando la tarea es compleja (Shiffrin y Schneider, 1977).

Al igual que Schneider y Shiffrin, Evans (1989) diferencia entre la memoria a corto y largo plazo. La memoria a corto plazo la asocia con guardar pequeñas cantidades de información que han sido recientemente procesadas y que se caracterizan por la rápida pérdida u olvido, salvo que sean trasladadas a la memoria a largo plazo. Evans afirma que, tradicionalmente, se ha identificado a la memoria a corto plazo con el procesamiento consciente (tipo 2) y con el alojamiento de diversos procesos cognitivos responsables del razonamiento. Del mismo modo, sostiene que la memoria a corto plazo actúa como una memoria intermedia entre la percepción y la memoria a largo plazo.

Smith y DeCoster (2000) también proponen dos sistemas de memoria asociados a cada tipo de procesamiento. Uno de ellos es el sistema de memoria de aprendizaje lento al que denominaron esquemático (*schematic*) y que se encarga de recordar información relacionada con la experiencia. Al segundo sistema de memoria lo denominaron de rápida vinculación (*fast binding*) y se asocia con el rápido aprendizaje de nueva información, por lo que, la nueva información puede ser recordada después de ocurrir de forma puntual. Los autores asocian el sistema de memoria de aprendizaje rápido con el procesamiento basado en reglas (procesamiento de tipo 2) mientras que el sistema de memoria de aprendizaje lento se asocia con el procesamiento asociativo (es decir, procesamiento de tipo 1). Además, el modelo sostiene que ambos sistemas de la memoria interaccionan entre sí a través del proceso de consolidación (*consolidation process*), en el que las representaciones aprendidas por el procesamiento basado en reglas son transferidas desde el sistema de aprendizaje rápido al sistema de aprendizaje lento, a través de la presentación repetida de dichas reglas. Tras el proceso de consolidación, la memoria deja de ser tan dependiente de la memoria de aprendizaje rápido para llevar a cabo la función de recuperación de la información. Sin embargo, cuando los sujetos no disponen de reglas relevantes para procesar el nuevo evento utilizan cualquier asociación que ellos tengan disponible; y, posteriormente, el sistema de memoria de aprendizaje rápido produce el conocimiento simbólico. De este modo, los dos sistemas de memoria intervienen en la consolidación de los recuerdos.

Brainerd y Reyna (2001) propusieron la teoría del Trazo Difuso (*Fuzzy-Trace Theory*) para explicar el procesamiento dual de la memoria y el razonamiento. Esta teoría, a diferencia de los MPD clásicos, se sustenta sobre la base de que se debe promover el razonamiento intuitivo sobre el razonamiento analítico. La teoría postula que, aunque, en general, la memoria y el razonamiento se encuentran ligados, hay situaciones en las que operan de manera disociada. Por ello, afirman que este modelo puede ser aplicado al funcionamiento de la memoria y del razonamiento de forma independiente. En relación con el modelo dual de razonamiento, los autores introducen dos tipos de razonamiento: el procesamiento del quiz de la cuestión (*gist processing*), el cual se identifica con el procesamiento de tipo 1, y el procesamiento difuso (*fuzzy processing*) más orientado a un razonamiento de tipo 2. En relación con el modelo dual de la memoria, esta teoría sostiene que existen dos tipos de memorias: las memorias del quiz de la cuestión (*gist memories*), más unida al procesamiento del quiz de la cuestión, y la memoria verbatim (*verbatim memories*) unida al procesamiento difuso. El funcionamiento de los dos tipos de memorias se sustenta en cuatro principios: (1) Los sujetos almacenan información específica del objeto (mediante un procesamiento difuso y es almacenadas en la memoria verbatim) e información general del objeto (mediante un procesamiento del quiz de la cuestión y es almacenada en la memoria del quiz de la cuestión). De este modo, ambos tipos de memoria

mantienen una relación paralela en su funcionamiento (*parallel verbatim-gist storage*). (2) Los resultados de los sujetos tienden a basarse especialmente en la recuperación de la información procedente de la memoria verbatim más que de la memoria del quiz de la cuestión (*dissociated verbatim-gist retrieval*). (3) Los recuerdos conscientes son el resultado de una recuperación de la memoria verbatim y los recuerdos globales son el resultado de la recuperación de la memoria del quiz de la cuestión (*memorial bases of conscious remembrance*). (4) los resultados obtenidos tienden a ser juicios de identidades categóricas (*identify and similarity judgments*).

Finalmente, el MPD desarrollado por Toates (2006) diferencia entre el procesamiento basado en estímulos (*stimulus-based process*) y el proceso de orden superior (*higher-order process*) para referirse a un procesamiento intuitivo o de tipo 1 y a un procesamiento consciente o de tipo 2, respectivamente. Toates indica que el funcionamiento del procesamiento basado en estímulos encierra representaciones del entorno que generan los mapas mentales o expectativas y desencadena el comportamiento en base a los estímulos percibidos y a la memoria implícita o procedimental. Sin embargo, el proceso de orden superior genera el contexto necesario en el que el proceso basado en estímulos opera y el resultado de ese procesamiento gana consciencia. Su funcionamiento se basa en la memoria explícita o declarativa.

2.3. Categorías y Atributos de los dos Tipos de Procesamiento

Debido a la gran variedad de modelos existentes, recientemente, se ha pretendido unificar los MPD desarrollados en las diversas disciplinas (Evans y Stanovich, 2013; Frankish, 2010). Para ello, se han propuesto *clústers* o categorías de propiedades asociadas a cada uno de los tipos de procesamiento. El estudio más amplio y reciente fue llevado a cabo por Evans y Stanovich (2013). Los autores identificaron cuatro *clústers*: consciencia, evolución, características funcionales y diferencias individuales. Sin embargo, no todos los investigadores han propuesto las mismas categorías de propiedades ni todos los MPD, en su formulación, han hecho referencia explícita a cada una de estas características. Por lo tanto, estos dos aspectos han hecho más difícil el esfuerzo de unificar los modelos.

En este apartado se presenta una nueva propuesta de *clústers* que trata de sintetizar y categorizar las principales propiedades de los MPD. La Tabla 2 enumera los *clústers* propuestos y las propiedades atribuidas a cada tipo de procesamiento en función del clúster. Los aspectos contemplados en esta tabla se explican en los siguientes apartados.

2.3.1. Características de los Procesamientos

Como se puede observar en la Tabla 2, el primer clúster o categoría de propiedades es el relacionado con las características que definen o identifican a los procesamientos.

Tradicionalmente, los investigadores han identificado las características de cada procesamiento cognitivo asociándolos con atributos opuestos. Así, algunos autores han definido los dos tipos de procesamiento como modos rivales de pensamiento (Hammond, 1996; Osman, 2004).

En términos generales, los diversos MPD presentan un alto consenso en cuanto a los atributos acuñados. Por ejemplo, el procesamiento de tipo 1 ha sido consistentemente definido como un **proceso intuitivo e impulsivo**, conducido, frecuentemente, por los factores viscerales, las intuiciones y percepciones del propio entorno (Brainerd y Reyna, 2001; Epstein, 2003; Kahneman y Frederick, 2002, 2005; Lieberman, 2003; Sloman, 1996; Strack y Deutsch, 2004). Sin embargo, el procesamiento de tipo 2 ha sido definido como un **proceso más analítico, racional y lógico** (Bargh y Chartrand, 1999; Bazerman et al., 1998; Epstein, 2003; Stanovich, 1999, 2011), dirigido por información abstracta y estadística y representaciones mentales que siguen los principios de jerarquía, lógica, causa-efecto y esquemas relacionales (Lieberman, 2003; Nisbett et al., 2001; Frankish, 2010).

Asimismo, el procesamiento de tipo 1 es identificado como un proceso automático, no intencional o involuntario y espontáneo, que no requiere de atención ni implica control cognitivo y, por ello, una vez iniciado es muy difícil detenerlo o ignorarlo. Sin embargo, el procesamiento de tipo 2 se asocia con una mayor deliberación y se define como intencional y voluntario, que requiere atención e implica control cognitivo, y por este motivo, el procesamiento analítico es más lento que el procesamiento automático (Chaiken, 1980; Chaiken et al., 1989; Guthrie et al., 2007; Haidt, 2001, 2008; Hammond, 1996; Metcalfe y Mischel, 1999; Smith y DeCoster, 2000; Stanovich y West, 2000; Toates, 2006).

Logan (1988), en su formulación de la teoría de la automatización, afirma que el procesamiento **automático** está construido sobre la adquisición de conocimiento procedente del entorno, que es almacenado en forma de representaciones y a los que denominó casos o ejemplos (*instances*). Igualmente, defiende que la automaticidad se adquiere mediante la experiencia consistente con el entorno y, por lo tanto, para considerar una acción automática es necesario que la decisión de actuar de ese modo haya sido directamente recuperada de la memoria. Cuando surge ese mismo problema recuperamos de la memoria el modo en que lo hemos resuelto previamente, automatizando el proceso y siendo más rápidos en nuestra actuación. En contraposición, cuando nos encontramos con un problema nuevo, donde no disponemos de experiencia previa y, por lo tanto, no podemos recuperar de la memoria la solución, el procesamiento no automático (procesamiento de tipo 2) será el que operará. No obstante, algunas de las funciones del procesamiento automático son desarrolladas de forma natural y no requieren de la consistencia de la experiencia para su desarrollo (Evans y Over, 1996; Bargh y Chartrand, 1999).

Tabla 2.

Categorías y Atributos de los dos Tipos de Procesamientos

Procesamiento de Tipo 1	Procesamiento de Tipo 2
Asociado con el Sistema 1	Asociado con el Sistema 2
Asociado con la memoria a LP/aprendizaje lento	Asociado con la memoria a CP/aprendizaje rápido
1- Características de los Procesamientos	
Intuitivo	Analítico
Impulsivo	Racional, lógico
Automático	Deliberado
No intencional, espontáneo	Intencional
Involuntario	Voluntario
No implica control cognitivo	Implica control cognitivo
No requiere de atención	Requiere de atención
Proceso no consciente o preconscious	Proceso consciente
Proceso rápido	Proceso más lento
Opera con poco o ningún esfuerzo	Opera con esfuerzo y concentración
Consume pocos recursos cognitivos	Demanda y consume recursos cognitivos
Alta capacidad cognitiva	Capacidad cognitiva limitada
Puede estar constantemente activo	No está constantemente activo
Eficiente	Eficaz y preciso
Poco flexible	Flexible
El proceso de aprendizaje y cambios en su aprendizaje es más lento	El proceso de aprendizaje y cambios en su aprendizaje es más rápido
Robusto: más resistente a interferencias y desórdenes clínicos	Sensible a interferencias y desórdenes clínicos
2- Elementos, Información y Principios mediante los que Operan	
Opera en base a la información sensorial y perceptual	Trata de describir el mundo a través del procesamiento de información abstracta
Los sentidos como fuente de conocimiento	Opera con información estadística y neutral
Base de las emociones, los miedos y las pasiones	Opera con complejas representaciones espaciotemporales y los pensamientos
Relacionado con las habilidades motoras	Habilidad para extraer pensamientos hipotéticos
No comprende el funcionamiento del procesamiento de tipo 2	Emocionalmente neutro
Opera mediante esquemas o representaciones mentales	Comprende el funcionamiento del procesamiento de tipo 1
Opera a través de impresiones y estereotipos	Opera basándose en reglas y normas
Lleva a cabo razonamientos inferidos	Relacionado con el razonamiento verbal e influido por las instrucciones verbales transmitidas culturalmente
Opera mediante heurísticos	El lenguaje y la cultura como fuentes de conocimiento
Opera mediante la adquisición de conocimiento aprendido de la relación y consistente experiencia con el entorno.	Lleva a cabo razonamientos deductivos
Opera mediante los principios de:	Puede operar con heurísticos
- Asociación	Genera juicios e inferencias a partir de información seleccionada por el procesamiento de tipo 1
- Codificación	Opera mediante:
- Sustitución	- Estructuras regladas
- Contigüidad	- Esquemas relacionales
Proceso concreto, contextualizado o de dominio específico	- Principio de causa-efecto
	- Simulación mental
	Proceso global, descontextualiza o de dominio general

Continúa

Tabla 2.

Continuación

Procesamiento de Tipo 1	Procesamiento de Tipo 2
3- Funciones de los Procesamientos	
Facilitar el trabajo de procesamiento de tipo 2	Inhibir la influencia del procesamiento de tipo 1
Llevar a cabo tareas rutinarias a través de la automatización del proceso	Controlar y supervisar el funcionamiento del proceso automático (corrección de los errores)
	Explicar y justificar su propio comportamiento (racionalización <i>post hoc</i>)
Complementar las limitaciones del proceso de tipo 2	Complementar las limitaciones del proceso de tipo 1
4- Arquitectura de los Procesamientos	
Procesamiento en paralelo	Procesamiento secuencial o lineal
5- Modulación de la Activación de los Tipos de Procesamiento	
<i>Variables Motivacionales</i>	
Precisión en el resultado	
Relevancia del mensaje	
Importancia percibida de las consecuencias	
Responsabilidad de la acción compartida o personal	
Objetivos individuales	
Importancia que el sujeto asigna a la tarea	
Regulación de las necesidades básicas	
Intensidad de la presencia de factores viscerales	
<i>Diferencias Individuales</i>	
Inteligencia	
Inteligencia emocional	
Disposiciones de pensamiento	
Capacidad de memoria operativa	
Capacidad de atención	
Edad	
Estados fisiológicos	
Daños cerebrales	
Consumo de fármacos	
Sexo	
<i>Variables Situacionales</i>	
Tiempo disponible para deliberar	
No disponer de juicios/información previa	
Características de la tarea	
Estrés	
6- Evolución Biológica	
Proceso evolutivamente más antiguo	Proceso evolutivamente más reciente
Proceso innato	Proceso
Proceso adaptativo (trata de adaptarse a los cambios del entorno)	Ayuda a adaptarnos a las nuevas necesidades de la vida moderna
Resultado de millones de años de evolución	Emerge del propio proceso basado en estímulos
Compartido con otros animales	Único del ser humano

Spunt (2015) define la **intencionalidad** como “el grado en el que el inicio del proceso depende de la presencia de una intención explícita de comenzarlo” (pág. 213). De acuerdo con esta definición, el procesamiento de tipo 2 se define como intencional, pues depende de la voluntad de activar y comenzar el proceso y su desarrollo requiere de atención. Sin embargo, el procesamiento de tipo 1 se define como involuntario o espontáneo, en el sentido de que no media una intención explícita de iniciar el proceso en sí mismo ni interviene la atención (Bargh y Chartrand, 1999; Haidt, 2001, 2008; Lieberman, 2003; Toates, 2006). El desempeño en pruebas de *stroop* y el efecto recencia (*priming effect*) demuestran que el procesamiento de tipo 1 es automático y que puede empezar y completar su proceso sin intención ni control cognitivo (Logan, 1988).

El rasgo de **control** hace referencia al “grado en que el curso del proceso puede ser alterado después de que haya sido iniciado” (Spunt, 2015, pág. 213). Toates (2006) relaciona el procesamiento de tipo 2 con la capacidad de controlar y dirigir nuestras acciones hacia la consecución de los objetivos e identifica la inhibición de la activación del proceso basado en estímulos (procesamiento de tipo 1) como una función que está bajo el control intencional del procesamiento de tipo 2. De este modo, se justifica la demanda de control cognitivo en el funcionamiento del procesamiento racional. Sin embargo, Logan (1988) apunta que, como resultado de la automatización, el procesamiento de tipo 1 es obligatorio, es decir, no se puede controlar y una vez iniciado es muy difícil detenerlo o ignorarlo.

Otro atributo comúnmente imputado a los procesamiento cognitivos es el relacionado con la **consciencia**. Spunt (2015) define este atributo como el “grado en que los sujetos son conocedores de los estímulos que inician el proceso, la operación del proceso en sí mismo y/o el resultado del proceso” (pág. 212). Diversos autores han postulado que el procesamiento de tipo 1 es no consciente, en el sentido de que apenas es conocedor del proceso en sí mismo, aunque sí pueda serlo del resultado (Glöckner y Witteman, 2010; Haidt, 2001, 2008; Smith y DeCoster, 2000; Strack y Deutsch, 2004; Wilson, Lindsey y Schooler, 2000). Otros, sin embargo, identifican al procesamiento de tipo 1 como preconscious (*preconscious* o *preattentive*) sugiriendo que no solo opera fuera de la consciencia, sino que también se configura como el paso previo al procesamiento consciente. Estos autores indican que la principal función del proceso preconscious es determinar qué información, de la analizada y codificada por el procesamiento automático, es enviada a la memoria operativa para posteriormente ser utilizada por el procesamiento analítico (Bargh y Chartrand, 1999; Evans, 1984, 2008; Smith y DeCoster, 2000; Stanovich, 2011). Por su parte, el procesamiento de tipo 2 ha sido considerado consciente por ser conocedor no solo del resultado de la actividad sino del proceso completo en sí mismo (Hammond, 1996; Haidt, 2001, 2008).

Los procesamiento cognitivos también están definidos por la **rapidez** de su funcionamiento. En este sentido, el procesamiento de tipo 1 se configura como un proceso rápido

en comparación con el procesamiento de tipo 2, comúnmente más identificado con un proceso lento (Evans y Over, 1996; Hammond, 1996; Metcalfe y Mischel, 1999; Pollock, 1989, 1991; Sloman, 1996). De acuerdo con la Teoría de las Instancias (*Instance Theory*; Logan, 1988), la rapidez del procesamiento de tipo 1 viene determinada por variables como el tiempo requerido para percibir el estímulo y emitir una respuesta, la práctica o experiencia previa con el evento y el tiempo necesario para aprender.

Otro rasgo atribuido a los tipos de procesamiento es el que está relacionado con la **capacidad cognitiva** que requieren para operar. El procesamiento de tipo 1 se suele identificar con la necesidad de invertir poco o ningún **esfuerzo** y con apenas consumir recursos cognitivos; mientras que el procesamiento de tipo 2 se asocia con la demanda y consumo de capacidad y recursos cognitivos y requiere de esfuerzo y concentración (Chaiken et al., 1989; Elqayam, 2009; Guthrie et al., 2007; Kahneman y Frederick, 2002, 2005; Stanovich, 1999, 2011). Así, el procesamiento de tipo 1 se define como un proceso de alta capacidad al poder estar constantemente activo y, en consecuencia, se ha sugerido que es el que tiende a operar por defecto (Elqayam, 2009; Reber, 1993). Sin embargo, el procesamiento de tipo 2 se define como un proceso que opera con esfuerzo y dispone de recursos cognitivos limitados, lo que provoca que el proceso consciente no se encuentre constantemente activo, sino que lo esté por pequeños periodos de tiempo. Además, se suele afirmar que el procesamiento de tipo 2 solo puede procesar una hipótesis o un modelo mental al mismo tiempo, es decir, opera de forma lineal o serial, siendo un indicador de que tiene una capacidad cognitiva más limitada (Evans y Over, 1996; Evans y Stanovich, 2013; Bargh y Chartrand, 1999).

En relación con esta característica, se encuentra el rasgo de eficiencia. El procesamiento de tipo 1 está, frecuentemente, relacionado con resultados **eficientes** (Pollock, 1989, 1991; Sloman, 1996). La eficiencia ocurre cuando el proceso puede ser ejecutado de forma rápida y en ausencia de atención (Spunt, 2015), esto es, alcanzar nuestro objetivo de procesamiento al menor coste o empleando la menor cantidad de recursos cognitivos. Sin embargo, el procesamiento de tipo 2, está más relacionado con resultados **eficaces y exhaustivos**, es decir, resultados más precisos o exactos, con menor tendencia al error o sesgo (Stanovich, 1999, 2011).

Asimismo, los procesos de tipo 2 han sido asociados con una mayor **flexibilidad** que los procesos de tipo 1 (Metcalfe y Mischel, 1999; Strack y Deutsch, 2004; Toates, 2006; Wilson et al., 2000). Strack y Deutsch (2004) afirman que el sistema impulsivo es más rígido en cuanto a la modificación de las representaciones y esquemas generados. Cambiar o modificar esa información y las relaciones establecidas entre ellas es un proceso muy lento. Wilson et al. (2000), en el modelo dual de actitudes propuesto, indican que las actitudes explícitas (derivadas del procesamiento de tipo 2) pueden cambiar o modificarse relativamente fácil, mientras que las actitudes implícitas (como los hábitos, derivadas del procesamiento de tipo 1) son más inflexibles, y su cambio es más lento. Del mismo modo, Pollock (1989, 1991) afirma que, al ser

extremadamente flexible, el procesamiento de tipo 2 puede gestionar cualquier tipo de situación. Por ello, asocia como principal función de la *intellección* (*intellection*; denominación acuñada por este autor al procesamiento de tipo 2), operar en aquellas situaciones en las que el razonamiento rápido e inflexible no puede actuar. La investigación ha sugerido que el proceso de aprendizaje del sistema 1 es lento y requiere de más tiempo para adquirir experiencia y aprender de la consistencia de la misma. Sin embargo, el aprendizaje y cambios en los conocimientos adquiridos por el sistema 2 son más rápidos y maleables (Epstein, 2003; Frankish, 2010; Smith y DeCoster, 2000).

Por último, otro rasgo que comúnmente se asocia con los procesamiento cognitivos es la **robustez** de los procesos, entendiendo por robustez su grado de vulnerabilidad en cuanto a la tendencia a sufrir interrupciones o interferencias durante su funcionamiento o a padecer daños cerebrales. A este respecto, Reber (1993) define la cognición implícita (procesamiento de tipo 1) como más robusta y resiliente que el proceso explícito (procesamiento de tipo 2) y menos propensa a sufrir interrupciones en su funcionamiento. Por lo tanto, es más resistente a las interferencias y a los desórdenes clínicos y, por ello, tiende a ser el proceso utilizado por defecto. Strack y Deutsch (2004) afirman que el sistema impulsivo es más sensible a la presencia de estímulos distractores, interfiriendo en su funcionamiento, por ello, opera más eficientemente ante niveles intermedios de arousal. Finalmente, Toates (2006) indica que el proceso de orden superior (procesamiento de tipo 2) es más susceptible a posibles perturbaciones y sus estructuras son más vulnerables que las del proceso basado en estímulos (procesamiento de tipo 1).

2.3.2. Elementos, Información y Principios Mediante los que Operan

El segundo clúster propuesto es el relacionado con los elementos, la información y los principios mediante los que opera cada tipo de procesamiento. Existe un amplio consenso entre los investigadores en afirmar que el *procesamiento de tipo 1* opera en base a la percepción, procesando **información sensorial y perceptual** o no verbal (dibujos e imágenes) que envuelven resultados motores. Por lo tanto, los sentidos son la fuente de conocimiento de este procesamiento (Hammond, 1996; Epstein, 2003; Lieberman, 2003; Smith y DeCoster, 2000; Strack y Deutsch, 2004). Asimismo, se asume que está relacionado con preferencias afectivas o emocionales y se considera la **base de las emociones, los miedos y las pasiones** (Bazerman et al., 1998; Epstein, 2003; Lieberman, 2003; Metcalfe y Mischel, 1999; Stanovich, 1999, 2011). Diversos autores afirman que el funcionamiento del procesamiento de tipo 1 está construido sobre la adquisición de **conocimiento procedente del entorno** y de nuestra experiencia con el mismo. Para ello, codifica la información que recibe del entorno y la experiencia, a través de las emociones y percepciones y crea **mapas o esquemas mentales** (por ejemplo, los estereotipos). Posteriormente, mediante la activación de la secuencia de esquemas mentales aprendidos se desencadena el

comportamiento específico (Epstein, 2003; Logan, 1988; Metcalfe y Mischel, 1999; Nisbett et al., 2001; Norman y Shallice, 1986).

Este tipo de procesamiento también se encarga, mediante **procesos asociativos**, de atribuir juicios a todos los eventos u objetos percibidos del entorno. De este modo, y de forma automática, explica y predice los acontecimientos y mensajes verbales procedentes del medio sobre la base de tales asociaciones (Kahneman y Frederick, 2002, 2005; Lieberman, 2003; Nisbett et al., 2001; Smith y DeCoster, 2000). Por eso se tiende a indicar que es un **proceso concreto, de dominio específico y que contextualiza automáticamente** los problemas a través de las intuiciones y percepciones (Elqayam, 2009; Evans, 2008; Nisbett et al., 2001; Smith y DeCoster, 2000; Stanovich y West, 2000). Hammond (1996) afirma que el procesamiento intuitivo depende de la recuperación de eventos muy específicos o vívidos y Evans y Over (1996) señalan que el sistema tácito (encargado del procesamiento de tipo 1) es también denominado **inferencial**, ya que tiende a aplicar conocimiento general a casos específicos.

Sin embargo, el *procesamiento de tipo 2* se ha identificado como emocionalmente neutro o **desprovisto de** todo tipo de **emociones** (Epstein, 2003; Metcalfe y Mischel, 1999). No se deja influir por las emociones, ya que se asocia con **información lógica, racional y abstracta** (Bazerman et al., 1998; Epstein, 2003; Frankish, 2010; Lieberman, 2003; Nisbett et al., 2001; Sloman, 1996) y, además, opera con complejas **representaciones espacio-temporales e información estadística y neutral** (Kahneman y Frederick, 2002, 2005; Metcalfe y Mischel, 1999). Asimismo, se ha señalado que este tipo de procesamiento depende de aspectos e **instrucciones verbales** que son culturalmente transmitidos (Evans y Over, 1996; Hammond, 1996; Lieberman, 2003; Stanovich, 1999, 2011). Por lo tanto, el lenguaje y la cultura son fuentes de conocimiento del procesamiento de tipo 2 (Epstein, 2003; Sloman, 1996).

Algunos autores, atribuyen a este tipo de procesamiento la capacidad para realizar **pensamientos hipotéticos** (Evans, 2006; Frankish, 2010; Stanovich, 1999, 2011) y afirman que es capaz de llevar a cabo un procesamiento de nivel más complejo que el procesamiento de tipo 1, lo que incluye hacer planificaciones, solucionar problemas, posponer las gratificaciones, realizar complejas generalizaciones y establecer relaciones de causa y efecto (Epstein, 2003; Toates, 2006). Además, el sistema racional es capaz de comprender el funcionamiento del sistema experiencial, permitiendo el entrenamiento de este y generando respuestas más apropiadas. Sin embargo, el sistema experiencial no es capaz de comprender el funcionamiento del sistema racional (Epstein, 2003).

Frecuentemente, se ha asumido que el procesamiento de tipo 2 está gobernado por **reglas o normas** que se crean a partir de la información seleccionada por el procesamiento de tipo 1 (Evans, 1984, 1989, 2006; Guthrie et al., 2007; Frankish, 2010; Kahneman y Frederick, 2002, 2005; Nisbett et al., 2001). Smith y DeCoster (2000) definen las reglas como “asociaciones explícitamente aprendidas de un evento singular que han sido simbólicamente representadas y

compartidas de forma interpersonal” (pág. 111). Sloman (1996) distingue diversos tipos de reglas. Algunas las define como simples instrucciones, otras se conforman sobre la base de las leyes, algunas son reglas descriptivas y, por último, están las reglas normativas. El resultado de un procesamiento basado en reglas suele reflejar mayor validez percibida al estar sustentado sobre reglas lógicas aprendidas y compartidas socialmente (Smith y DeCoster, 2000). Kruglanski y Gigerenzer (2011) indican que las reglas pueden ser adquiridas a través de la experiencia personal, a través del desarrollo social o a través de la cultura; y las definen como potencialmente maleables, en el sentido de que pueden ser aprendidas y desaprendidas, olvidadas y recuperadas de la memoria. Además, Nisbett y colaboradores (2001) defienden que el procesamiento de tipo 2 subyace de la descontextualización del contenido para aplicar la lógica abstracta y de evitar las contradicciones. Por lo tanto, se define como un **proceso de dominio general**, enfocado a información muy concreta y que sirve para **descontextualizar y despersonalizar los problemas**, es decir, procesa problemas sin contenido social (Elqayam, 2009; Evans, 2008; Evans y Over, 1996; Nisbett et al., 2001; Smith y DeCoster, 2000; Stanovich y West, 2000).

En relación con los dos tipos de procesamiento, se han llevado a cabo varios estudios para explorar la relación entre los principales heurísticos y el razonamiento humano. Epstein (2003), de acuerdo con Tversky y Kahneman (1974), ha definido los heurísticos como “atajos cognitivos que la gente emplea espontáneamente en la toma de decisiones en condiciones de incertidumbre y que son una fuente prevalente de reacciones irracionales en multitud de situaciones” (pág. 166). Anderson, Cullen y Stamoulis (2008) los definen como simples reglas que permiten tomar decisiones ante la ausencia de la totalidad de información o cuando se es incapaz o no se está dispuesto a incorporar toda la información disponible en la toma de decisión. Tradicionalmente, se ha tendido a relacionar el uso de **heurísticos cognitivos** con el procesamiento de tipo 1, ya que se cree que, al operar fuera de la consciencia, sin atención ni control y de forma más rápida que el procesamiento de tipo 2, puede sesgar en mayor medida sus respuestas. No obstante, varios autores han señalado la posibilidad de que el procesamiento de tipo 2 también pueda estar sesgado por los heurísticos y atajos cognitivos (Brainerd y Reyna, 2001; Chaiken et al., 1989; Guthrie et al., 2007; Hammond, 1996; Stanovich, 1999, 2011). Kahneman y Frederick (2002, 2005) asumen que los errores y sesgos cognitivos únicamente ocurren cuando ambos tipos de procesamiento fallan en su funcionamiento. Evans (2006) señala que, aunque la tendencia del procesamiento analítico es a inhibir los sesgos emitidos por el procesamiento heurístico, en ocasiones el proceso de tipo 2 es propenso a sesgar su propio proceso, debido, especialmente, al principio de satisfacción (*satisficing principle*). Este principio suscribe que cada representación es evaluada por el procesamiento de tipo 1 o de tipo 2, en relación con nuestros objetivos, y aceptada satisfactoriamente por nuestra mente. En la creación de pensamientos, si el sistema analítico trabaja con las representaciones heurísticas elaboradas por el procesamiento de tipo 1, puede componer respuestas sustentadas sobre inferencias falaces. De este modo, el procesamiento de

tipo 2 estaría sesgando sus respuestas. En el capítulo 3 de esta tesis se recogen y definen los sesgos cognitivos más relevantes desde el punto de vista de la prevalencia.

Por último, los MPD han manifestado que cada uno de los tipos de procesamiento conectan sus elementos de información a través de diferentes *tipos de relaciones* (Strack y Deutsch, 2004). El procesamiento de tipo 1 está vinculado a los **principios de asociación, similitud, contigüidad y categorización** (Smith y DeCoster, 2000). Sloman (1996) indica que el sistema asociativo lleva a cabo inferencias y predicciones sobre la base de los principios de similitud y de contigüidad temporal. Nisbett et al. (2001) identifican el procesamiento holístico como asociativo y afirman que refleja similitud y contigüidad cuando opera. Strack y Deutsch (2004) señalan que la información del sistema impulsivo está almacenada a través de representaciones asociativas y categóricas, de modo que, cuando un elemento es activado, esta activación se difunde hacia otros elementos en función de la frecuencia y la recencia de previas activaciones.

El procesamiento de tipo 2 está más asociado con **estructuras regladas**, principios relacionales y de **causa-efecto**, así como **principios de simulación o suposición mental**. Nisbett et al. (2001) indican que el proceso analítico se centra en sistemas simbólicos de representaciones y su funcionamiento refleja estructuras regladas. Strack y Deutsch (2004) proponen que la información del sistema racional se almacena en forma de representaciones que conectan un elemento con otro, de acuerdo con esquemas relacionales, y asignan un valor a la relación. Evans (2008; Frankish, 2010; Stanovich, 2011) sostiene que el procesamiento del sistema 2 incluye la habilidad para extraer pensamientos hipotéticos por vía de la simulación y de las suposiciones mentales. Finalmente, Sloman (1996) señala que el sistema basado en reglas trata de describir el mundo capturando diferentes tipos de estructuras lógicas y causales bien organizadas.

Por su parte, Hammond (1996) defiende dos *principios de funcionamiento* mediante los que operan los dos tipos de procesamiento. Concretamente indica que la cognición humana es capaz de emplear patrones de reconocimiento y relaciones funcionales cuando opera. Sostiene que las relaciones funcionales hacen referencia a que la cognición busca, de forma alternativa, explicaciones basadas en el **principio de coherencia o en el principio de correspondencia** (Haidt, 2001, 2008; Sloman, 1996). El patrón de reconocimiento, sin embargo, hace referencia a que las personas responden de acuerdo con modelos, es decir, los sujetos juzgan situaciones en base a otras situaciones que consideran similares. Finalmente, el autor defiende que ambos heurísticos pueden sesgar los juicios y la toma de decisión de los sujetos.

En resumen, el procesamiento de tipo 1 es un proceso que opera en base a información sensorial y emocional percibida y aprendida de su relación con el entorno. Crea representaciones y esquemas mentales (por ejemplo, estereotipos) con toda la información recogida del medio y asocia, juzga y contextualiza los estímulos percibidos de acuerdo con esa información. Para ello, emplea relaciones y principios de asociación, categorización y similitud y contigüidad. Sin

embargo, el procesamiento de tipo 2 es un proceso lógico y racional que opera con información abstracta, estadística y emocionalmente neutra. Está relacionado con el razonamiento verbal e hipotético y su funcionamiento se basa en la creación de reglas y normas transmitidas culturalmente. Su tendencia es a descontextualizar los problemas y procesarlos sin contenido social. Organiza la información en base a estructuras regladas y los principios relacionales, de causa-efecto y de simulación o suposición mental. Aunque el procesamiento de tipo 1 es más propenso a sesgar sus respuestas en base a la aplicación de atajos cognitivos, el procesamiento de tipo 2 también puede operar empleando heurísticos.

2.3.3. *Funciones de los Procesamientos*

Las principales funciones que se han atribuido al *procesamiento de tipo 1* han sido la **percepción y codificación de la información** recibida del entorno, con dos principales objetivos: (1) facilitar el trabajo del procesamiento de tipo 2, al proveer contenido a la memoria operativa y determinar qué información entra en el procesamiento analítico (Evans, 1984, 1989, 2008) y (2) se encarga de **actividades rutinarias** a través de la automatización del proceso mediante el aprendizaje y la consistente experiencia con el entorno (Bargh y Chartrand, 1999; Logan, 1988; Pollock, 1989, 1991; Smith y DeCoster, 2000).

Las funciones que tradicionalmente se han asociado al *procesamiento de tipo 2* son más numerosas y variadas. Por ejemplo, una de las características es que implica control cognitivo en su funcionamiento. El procesamiento de tipo 2 es capaz de controlar y dirigir nuestras acciones hacia la consecución de los objetivos, generando inferencias y juicios y resolviendo problemas. Para ello, en ocasiones, debe inhibir la activación del proceso de tipo 1 para que no interfiera en su funcionamiento (Toates, 2006). En este sentido, una de las principales funciones atribuidas al procesamiento de tipo 2 es **inhibir, anular o ignorar los resultados del procesamiento de tipo 1** cuando sea necesario (Alter, Oppenheimer, Epley y Eyre, 2007; Evans, 2008; Pollock, 1989, 1991; Kahneman y Frederick, 2002, 2005; Toates, 2006; Guthrie et al., 2007). Stanovich (1999, 2011) indica que, para llevar a cabo esta función, el procesamiento de tipo 2 requiere de dos habilidades: una que le permita interrumpir el proceso cognitivo automático y otra que le permita suprimir la tendencia de la respuesta. Para poder suprimir la respuesta es necesario que el procesamiento de tipo 2 ofrezca una respuesta alternativa. Esta respuesta es el resultado del razonamiento hipotético o de simulaciones mentales que temporalmente crea nuestra mente para testar o comprobar las representaciones del mundo real. Metcalfe y Mischel (1999) identifican el autocontrol y la autorregulación (control de los impulsos) como un ejemplo de la función de inhibición del procesamiento de tipo 1 por parte del procesamiento de tipo 2 (Bechara, 2004, 2005; Loewenstein, 1996).

Sloman (1996) señala que el procesamiento de tipo 2 es responsable de **controlar y supervisar el procesamiento del sistema 1**, es decir, se encarga de corregir y/o anular el

procesamiento intuitivo cuando falla (Epstein, 2003; Hammond, 1996; Kahneman y Frederick, 2002, 2005). Epstein (2003) indica que el procesamiento lento del sistema racional permite corregir el procesamiento del sistema experiencial, al reconocer lo inapropiado de sus pensamientos impulsivos y generar unos más constructivos. Kahneman y Frederick (2002, 2005) sugieren que el sistema 1 propone de forma inmediata respuestas intuitivas para juzgar los problemas y que el sistema 2 comprueba la calidad de las propuestas respaldando su adecuación o anulándolas para dar paso a la elaboración de una respuesta más sofisticada y analítica. Por lo tanto, el procesamiento de tipo 2 no solo se encarga de la inhibición de la respuesta del proceso de tipo 1, sino que también lo supervisa para poder detectar y corregir posibles errores o sesgos.

Otra de las funciones habitualmente asignadas al procesamiento de tipo 2 es la función de **racionalización de las percepciones y los resultados intuitivos** obtenidos por la aplicación del procesamiento de tipo 1. Es decir, los sujetos tienden a tratar de explicar y justificar su propio comportamiento a través del conocimiento de la situación. Sin embargo, esta racionalización tiende a ser *post hoc*, es decir, después de llevar a cabo el comportamiento, los sujetos razonan sobre su justificación y causas (Epstein, 2003; Evans y Wason, 1976; Frankish, 2010; Haidt, 2001, 2008; Strack y Deutsch, 2004; Wason y Evans, 1975). En este sentido, Bazerman et al. (1998) encontraron que, cuando se siente la necesidad de justificar la decisión tomada, la tendencia es a que opere el procesamiento de tipo 2, ya que la justificación nos obliga a analizar en profundidad la información y las opciones.

Varios autores han identificado que el proceso de tipo 2 **complementa el procesamiento de tipo 1**. Pollock (1989, 1991) indica que el procesamiento racional opera en aquellas situaciones en las que el razonamiento rápido e inflexible no puede actuar. Por ejemplo: (1) situaciones novedosas donde el proceso de tipo 1 no dispone de una secuencia de actuación aprendida previamente (Lieberman, 2003; Logan, 1988; Norman y Shallice, 1986; Toates, 2006); (2) situaciones en las que el sujeto desee anticiparse al futuro (Evans y Over, 1996) y (3) situaciones complejas en las que el procesamiento de tipo 1 falla a la hora de operar (Alter et al., 2007; Bazerman et al., 1998; Hammond, 1996). En todos estos casos, el procesamiento de tipo 2, al ser más lento, es capaz de proveer una solución adecuada (Toates, 2006).

No obstante, algunos MPD admiten que el procesamiento de tipo 1 también puede **complementar el procesamiento de tipo 2**, operando en aquellas situaciones en las que el razonamiento analítico no puede (Bazerman et al., 1998). Así, por ejemplo, Conrey et al. (2005) afirma que un ejemplo de proceso automático es aquel encaminado a facilitar una respuesta cuando el procesamiento controlado no puede operar o falla en su intento de operar, aunque su funcionamiento implique sesgos y heurísticos. Por su parte, Bazerman y colaboradores (1998) defienden que, ante situaciones de ambigüedad, el procesamiento de tipo 2 no puede funcionar y será el proceso de tipo 1 el que debe asumir el control de la respuesta.

2.3.4. *Arquitectura de los Procesamientos*

Tradicionalmente, se han diferenciado dos tipos de funcionamiento de los procesos cognitivos: el funcionamiento en paralelo y el funcionamiento serial o secuencial. Neisser (1963) identifica al procesamiento en paralelo como un procesamiento múltiple (*multiple processing*), donde las operaciones no son realizadas literalmente de forma simultánea, ya que esto resultaría imposible por las propias limitaciones físicas de nuestra mente. Se dice que su funcionamiento es simultáneo en el sentido de que todos los procesos que lo conforman son completados antes de tomar una decisión. Sin embargo, el procesamiento secuencial es aquel donde la toma de decisión podría preceder a otra toma de decisión, es decir, la segunda pregunta depende de la respuesta dada a la primera pregunta; y, por lo tanto, no se pueden llevar a cabo diversas acciones de forma simultánea. El autor indica que dos actividades podrán ser desempeñadas simultáneamente si: (1) ambas tareas no requieren una respuesta; (2) la realización de ambas operaciones no envuelve un elevado nivel de complejidad; y (3) al menos una de las actividades se puede llevar a cabo de forma automática, sin requerir un procesamiento consciente.

El procesamiento de tipo 1 ha sido asociado con un **funcionamiento en paralelo**, simultáneamente con otro proceso automático o con un proceso más controlado. Sin embargo, el procesamiento de tipo 2 ha sido relacionado con un **funcionamiento serial o secuencial**, es decir, lineal o sistemático, que opera por pasos. La asociación del procesamiento de tipo 1 con un proceso en paralelo se justifica en base a las características de rapidez en el procesamiento y a la automaticidad del proceso. Sin embargo, la asociación del procesamiento de tipo 2 con un proceso lineal se ha justificado en base al lento procesamiento y a la consciencia del sistema 2 (Evans, 2008; Evans y Over, 1996; Evans y Stanovich, 2013; Lieberman, 2003; Pollock, 1989, 1991; Sloman, 1996).

Cuando se ha tratado de establecer la relación funcional existente entre los dos tipos de procesamiento han surgido discrepancias entre los autores. Smith y DeCoster (2000) afirman que existe un debate acerca de cómo es el proceso de interacción entre ambos sistemas. Los autores estipulan que se ha tratado de diferenciar entre procesamientos alternativos (es decir, opera un sistema o el otro, pero no ambos), procesamientos secuenciales (primero opera el sistema automático por defecto y después el sistema basado en reglas puede continuar si lo desea), o si los procesamientos son simultáneos (es decir, operan ambos coetáneamente).

Evans (2008) y Evans y Stanovich (2013) han tratado de clasificar los MPD en dos grandes grupos en función del modo de conexión o relación entre los sistemas o procesamientos. El primer grupo está formado por las teorías por defecto-intervencionistas (*default-interventionist*) que postulan que el razonamiento analítico y deliberado puede ser aplicado cuando pretendemos inhibir la respuesta intuitiva y sesgada del sistema 1 y reemplazarla por una respuesta basada en el razonamiento analítico. Esta teoría defiende que el procesamiento de tipo 1 opera por defecto (está siempre activo) y el procesamiento de tipo 2 se activará voluntariamente

cuando se desee inhibir la respuesta del procesamiento de tipo 1. Un ejemplo de este tipo de teorías sería el MPD de Chaiken et al. (1989). Estos autores proponen que los sujetos tendemos a procesar con el menor esfuerzo posible, somos ahorradores en términos del esfuerzo cognitivo. Este principio está relacionado con el concepto de “misericordia cognitiva” (*cognitive miser*; Simon, 1955; Toplak et al., 2014; Tversky y Kahneman, 1974), y se refiere a la tendencia básica de los humanos a emplear el procesamiento de tipo 1 por ser más rápido, requerir menor esfuerzo cognitivo y consumir menos energía mental. El MPD propuesto por Stanovich (1999) también se asienta sobre este principio. Otros ejemplos de modelos por defecto-intervencionistas serían la teoría del razonamiento heurístico-analítico postulada por Evans (2006), el modelo dual de Haidt (2001, 2008) o la teoría del trazo difuso de Brainerd y Reyna (2001).

El segundo grupo de teorías son las teorías paralelo-competitivo (*parallel-competitive*) que defienden que ambos sistemas operan en paralelo y compiten por dominar la respuesta final. Este tipo de teorías se basan en la idea de que existen dos formas de aprendizaje, dando lugar a dos tipos de conocimientos (implícito y explícito) que posteriormente tratarán de competir por el control del comportamiento. Ejemplos de MPD que apoyan esta hipótesis serían los modelos propuestos por Sloman (1996), Smith y DeCoster (2000) y Kahneman y Frederick (2002).

No obstante, Conrey et al. (2005) han sugerido que la diferenciación hecha por Evans (2008) entre modelos duales paralelos-competitivos y los modelos por defecto-intervencionista pueden sumarse en un único modelo que simultáneamente produzca ambos tipos de interacción entre los procesamientos. Los autores encasillan el modelo que proponen dentro de esta tercera categoría.

Por su parte, Elqayam (2009) presenta un modelo bidimensional de las diferencias arquitectónicas o de configuración de los MPD de la información. El autor denominó a la primera *dimensión temporal-secuencial* y se refiere a la diferenciación propuesta por Evans (2008), en cuanto a que el procesamiento dicotómico funciona (1) en paralelo y compiten por el control de la respuesta (*paralelo-competitivo*) o, por el contrario, (2) funciona de forma secuencial, iniciado por el proceso heurístico (*por defecto-intervencionista*). La segunda dimensión fue propuesta por Elqayam (2009) y la denominó *dimensión funcional*. Esta dimensión diferencia entre modelos duales en función de la relación funcional existente entre los procesos cognitivos. Así, señala los *modelos interaccionistas*, donde el proceso heurístico afecta al proceso analítico y/o viceversa y los *modelos independentistas* donde los dos tipos de procesos operan de forma independiente y no influyen el uno sobre el otro. Ejemplos de modelos interaccionistas serían, por ejemplo, los MPD propuestos por Evans (2006), Stanovich (1999, 2011) o Guthrie et al. (2007). Por su parte, serían ejemplos de modelos independentistas el propuesto por Chaiken et al. (1989) o el de Smith y DeCoster (2000).

La nueva dimensión de Elqayam (2009) admite que los modelos paralelos-competitivos y los modelos por defecto-intervencionistas pueden ser tanto modelos independientes como

interaccionistas. Así, los MPD desarrollado por Kahneman y Frederick (2002, 2005) y el modelo de Smith y DeCoster (2000) son definidos como modelos paralelos-competitivos, sin embargo, el primero es interaccionista ya que, el sistema 1 puede afectar al procesamiento del sistema 2, mientras que, el segundo es independentista. Por su parte, los modelos de Stanovich (1999, 2011) y el de Chaiken et al. (1989) son clasificados como modelos por defecto-intervencionista, sin embargo, el primero es considerado interaccionista y el segundo puede ser definido como independentista de acuerdo con la hipótesis aditiva. La hipótesis aditiva indica que ambos tipos de procesamiento pueden ejercer efectos aditivos, es decir, operar de forma independiente. No obstante, el MPD desarrollado por Chaiken y colaboradores (1989) puede ser también clasificado como interaccionista si atendemos a la hipótesis del sesgo. De acuerdo con esta hipótesis, el procesamiento heurístico puede sesgar al procesamiento sistemático y, por lo tanto, la relación funcional sería interaccionista.

A continuación, se presenta una clasificación de los MPD, de acuerdo con la estructura bidimensional propuesta por Elqayam (2009; Evans, 2008). Únicamente se han clasificado los modelos que, en su formulación, han presentado información que permitiese la categorización. La Tabla 3 recoge el resultado de la codificación.

Tabla 3.

Clasificación de los MPD de Acuerdo con la Arquitectura Propuesta por Elqayam (2009) y Evans (2008)

Autores	Dimensión temporal-secuencial		Dimensión funcional	
	Por defecto-intervencionista	Paralelo-competitivo	Interaccionista	Independentista
Thaler y Shefrin (1981)			X	
Norman y Shallice (1986)	X		X	
Logan (1988)	X		X	
Chaiken et al. (1989)	X		X ¹	X ²
Reber (1993)		X		
Sloman (1996)		X	X	
Metcalf y Mischel (1999)		X	X	
Stanovich (1999, 2011)	X		X	
Smith y DeCoster (2000)		X		X
Wilson et al. (2000)	X		X	
Brainerd y Reyna (2001)	X			
Haidt (2001, 2008)	X		X	
Kahneman y Frederick (2002, 2005)		X	X	
Epstein (2003)		X	X	
Strack y Deutsch (2004)		X	X	
Evans (2006)	X		X	
Toates (2006)		X	X	
Guthrie et al. (2007)		X	X	

Nota. ¹Clasificado como interaccionista de acuerdo con la hipótesis del sesgo. ²Clasificado como independentista de acuerdo con la hipótesis aditiva.

2.3.5. *Modulación de la Activación de los Procesamientos*

Algunos MPD postulan que la activación de los procesamientos cognitivos se encuentra enfatizada o moderada por diversos factores. Las principales variables examinadas al respecto han sido clasificadas como factores motivacionales, diferencias individuales o factores situacionales. Frecuentemente, los MPD incluyen en sus teorías varias de esas variables.

2.3.5.1. *Variables Motivacionales*

Smith y DeCoster (2000) afirman que la motivación es un aspecto relevante para que el sistema basado en reglas opere. Si el sujeto no está lo suficientemente motivado será el sistema asociativo el que tenderá a operar (Wilson et al., 2000). Señalan que, en ocasiones, la motivación es simplemente alcanzar la **precisión** deseada en el resultado (Haidt, 2001, 2008; Kruglanski y Gigerenzer, 2011; Smith y DeCoster, 2000). Chaiken y colaboradores (1989) sostienen que para activar un proceso sistemático (es decir, encargado del procesamiento de tipo 2) es necesaria la presencia de factores motivacionales. Además, indican que los principales factores motivacionales que intervienen están relacionados con la **relevancia del mensaje** para el receptor, la importancia percibida de las **consecuencias** que se puedan desencadenar de los juicios hechos, si la **responsabilidad** de la evaluación del mensaje es de ese receptor o compartida con alguien más y los **objetivos individuales** de procesamiento. Los autores han mostrado que los sujetos con alta motivación y, por ende, con mayor capacidad para procesar racionalmente, son menos propensos a que los heurísticos (procesamiento de tipo 1) produzcan un impacto persuasivo en su actitud hacia el mensaje (hipótesis de atenuación). Kruglanski y Gigerenzer (2011) añaden el grado de **importancia** que el sujeto asigna a la **tarea** como un factor motivacional adicional.

Strack y Deutsch (2004) contemplan la **regulación de las necesidades básicas** como otro mediador de la activación de los procesos cognitivos. Señalan que, si las necesidades básicas de las personas (tales como comer, beber, dormir, etc.) no son saciadas, se llevarán a cabo, de forma automática e incontrolada, comportamientos y acciones específicos con la finalidad de remediar esa deficiencia. Por lo tanto, el sistema automático se activa más fácilmente ante la presencia de necesidades básicas no satisfechas. Loewenstein (1996) postula que el control de los impulsos requiere manipular o anular la **intensidad de la presencia de los factores viscerales** (definidos como antojos asociados con los impulsos). Este autor indica que, a niveles suficientemente altos de intensidad, los sujetos centran su atención y motivación en actividades relacionadas con los factores viscerales, es decir, con saciar los antojos impulsivos. Mientras que, a niveles bajos de intensidad, los sujetos son capaces de controlar cognitivamente los factores viscerales de una forma óptima.

2.3.5.2. *Diferencias Individuales*

Varios autores han sugerido que las diferencias individuales influyen en la activación de los procesos cognitivos. Entre las diferencias individuales estudiadas se encuentran la capacidad cognitiva, la capacidad de la memoria operativa, la capacidad de atención, la inteligencia emocional, las disposiciones de pensamiento o estilos cognitivos, la edad y el sexo.

Por ejemplo, Evans (2006), a través de la teoría del razonamiento heurístico-analítico, ha señalado la **inteligencia** y la **memoria operativa** como dos variables individuales moduladoras del funcionamiento de los tipos de procesamiento. Varios MPD han afirmado la existencia de una correlación entre el procesamiento de tipo 2 y la capacidad mental general y/o la capacidad de la memoria operativa (Kruglanski y Gigerenzer, 2011; Evans y Stanovich, 2013; Epstein, 2003; Stanovich, 1999, 2011; Wilson et al., 2000). De modo que personas altas en capacidad mental general o memoria operativa tenderán a inhibir mejor el procesamiento de tipo 1 y a obtener un mejor resultado, comportamiento o razonamiento (Evans, 2006; Haidt, 2001, 2008; Stanovich y West, 2000). Metcalfe y Mischel (1999), en los estudios llevados a cabo sobre el aplazamiento de la gratificación con niños de preescolar, mostraron que el número de segundos que los menores son capaces de esperar por obtener una gratificación mayor, en comparación con la inmediata recompensa, predice, décadas después, resultados sociales y cognitivos (como, por ejemplo, los resultados en el SAT). Sin embargo, la evidencia empírica sugiere que el procesamiento de tipo 1 es independiente de la capacidad mental general (Evans y Stanovich, 2013; Reber, 1993).

Otros autores (Epstein, 2003; Stanovich y West, 2000) han asociado la **inteligencia emocional** al procesamiento de tipo 1. Epstein (2003) afirma que la inteligencia emocional mide aspectos más ligados con la personalidad, como la motivación, el ego, las emociones, las habilidades sociales y la creatividad que se asocian de forma más robusta con el procesamiento de tipo 1.

La moderación de la activación del procesamiento de tipo 2 también se puede dar a través de factores que afectan a la capacidad cognitiva (Smith y DeCoster, 2000). Por ejemplo, emplear **estímulos distractores** tanto externos como internos o utilizar el tiempo como presión. En el primer caso, como el sistema basado en reglas opera de forma más lenta y requiere de mayor carga cognitiva o atención, es más probable que este procesamiento se vea interrumpido por distractores, interferencias o cualquier otro elemento. En relación con este punto, Metcalfe y Mischel (1999), dentro del ámbito de estudio del retraso en las gratificaciones, han enumerado diversas situaciones en las cuales uno de los sistemas domina al otro. Por ejemplo: (1) presentar de forma oculta o vagamente el estímulo que activa el funcionamiento del procesamiento de tipo 1 cuando los sujetos están tratando de inhibir su respuesta impulsiva, facilita el retraso de la gratificación. (2) La fuerza del estímulo que activa el funcionamiento del sistema 1 puede disminuir si el estímulo es presentado mientras el sujeto está tratando de inhibir la respuesta impulsiva, ya que se evita o se reduce la atención al estímulo (a esto se le denomina autocontrol

interno). (3) Cuando se presenta el estímulo que activa el sistema 1, la fuerza de este estímulo decrece con la exposición del sujeto a un segundo estímulo externo, denominado distractor, que activa el sistema 2 o el propio sistema 1, pero distrayéndolo del objeto diana. Asimismo, la fuerza del estímulo puede ser disminuido generando internamente auto-distractores. (4) Cuando se presenta información de un estímulo o su imagen en lugar del propio objeto en sí, el sistema 2 se activa e inhibe el procesamiento del sistema 1. (5) Por último, cuando una persona piensa en las propiedades “frías” y aspectos del estímulo más que en las propiedades “cálidas” del mismo, el sujeto consigue retrasar el comportamiento y/o la recompensa, controla sus impulsos.

En relación con la **atención** a los estímulos, Kruglanski y Gigerenzer (2011) indicaron que la capacidad de atención afecta a la aplicación de reglas y, por lo tanto, a la activación de uno u otro proceso cognitivo. Los autores afirman que la capacidad de atención depende de la carga cognitiva que la tarea exija así, como el grado de fatiga mental del sujeto. Strack y Deutsch (2004) asumen que la activación del procesamiento de tipo 2 depende de la intensidad y atención en el procesamiento del estímulo. De modo que el sistema impulsivo (procesamiento de tipo 1) procesa la información y, dependiendo de la intensidad y la atención con que se recibe, el estímulo podrá ser también procesado por el sistema reflexivo (procesamiento de tipo 2). Por lo que ambos sistemas procesarán en paralelo. Finalmente, Logan (1988) afirma que la recuperación de la información almacenada en la memoria es una consecuencia obligatoria e inevitable de procesar mediante la atención, por lo tanto, la cantidad y calidad de la atención influye en la capacidad del procesamiento de tipo 2 para recuperar la información.

Por otro lado, Evans (2006) y Stanovich (2011) indican que otras variables disposicionales, como los **estilos cognitivos**, determinan la activación e intervención del procesamiento analítico. Stanovich (1999) define las disposiciones de pensamiento como “mecanismos y estrategias psicológicas relativamente estables que generan tendencias y tácticas comportamentales características” (pág. 36). Stanovich (2011) afirma que, aunque la inteligencia correlaciona con tareas de pensamiento racional, no son una medida exhaustiva del funcionamiento cognitivo pues fallan a la hora de evaluar importantes estrategias y estilos cognitivos que son componentes críticos de una mente reflexiva. La principal diferencia entre inteligencia y disposiciones de pensamiento (por ejemplo, la escala *Need for Cognition*) es que el primero indica lo que la gente puede hacer y el segundo lo que está dispuesta o inclinada a hacer (Evans y Stanovich, 2013). Sternberg (1988) señala que los estilos intelectuales representan una importante unión entre la inteligencia y la personalidad. Probablemente representan una forma a través de la cual, la personalidad se manifiesta en los pensamientos inteligentes y la conducta. En relación con esto, diversos autores han mostrado una relación entre el procesamiento de tipo 2 y los estilos cognitivos analíticos (Chaiken et al., 1989; Juanchich, Dewberry, Sirota y Narendran, 2016; Leykin y DeRubeis, 2010).

La tercera diferencia individual que ha sido propuesta como variable moduladora de la activación de los procesos cognitivos ha sido la **edad**. En este sentido, se ha afirmado que el procesamiento de tipo 1 es independiente de la edad del sujeto (Reber, 1993; Evans y Over, 1996). Stanovich (2011) afirma que los adolescentes son cognitivamente más sofisticados que los niños y los adultos más cognitivamente sofisticados que los adolescentes. Por su parte, Toates (2006) identifica una función invertida en forma de U para referirse a la relación existente entre la edad y la activación del procesamiento de tipo 2. Afirma que, con el desarrollo, se incrementa la eficacia del sistema 2 para activar el procesamiento analítico, seguida por una disminución de su eficacia con el incremento de la edad. Del mismo modo, Metcalfe y Mischel (1999) sostienen que el sistema “cálido” se desarrolla antes que el sistema “frío”, por lo que, en las primeras fases de desarrollo de la vida, el sistema “frío” no está totalmente presente y las conexiones entre los nodos de ambos sistemas aún no han sido establecidas. En la temprana edad, el sistema “cálido” domina al sistema “frío” y el control de los impulsos y el retraso en las gratificaciones predominan con la edad.

Por último, se deben enumerar otras variables individuales que han sido menos examinadas, pero igualmente consideradas variables moduladoras de la activación de los procesos. Por ejemplo, Toates (2006) integra en su modelo los **estados fisiológicos** de los sujetos, es decir, asume que los niveles hormonales y los nutrientes del cuerpo influyen en el procesamiento de tipo 2. Del mismo modo, el consumo de diversos **fármacos** o la aparición de **daños cerebrales** puede fomentar o minar el funcionamiento del procesamiento de tipo 1 o del procesamiento de tipo 2 (Metcalfe y Mischel, 1999; Toates, 2006). Es más probable que ante cualquiera de esas dos situaciones se active el procesamiento de tipo 1, ya que este tipo de procesamiento es menos vulnerable a sufrir interrupciones o interferencias durante su funcionamiento o a padecer daños cerebrales (Reber, 1993; Strack y Deutsch, 2004; Toates, 2006).

Finalmente, Epstein (2003) sugiere que la variable **sexo** también puede influir en la tendencia a activar un tipo de procesamiento u otro. Afirma que las mujeres son más propensas a aplicar el procesamiento de tipo 1 que los hombres, siendo estos los que se inclinan hacia un mayor *engagement* en el procesamiento racional. Sin embargo, señala que las diferencias suelen ser pequeñas.

2.3.5.3. Variables Situacionales

Las variables situacionales también han sido señaladas como posibles moderadores del funcionamiento y activación de los tipos de procesamiento. En primer lugar, una de las variables situacionales más citadas ha sido la necesidad de disponer de **tiempo** para deliberar y evitar, así, una respuesta impulsiva o sesgada (Evans, 2006; Haidt, 2001, 2008; Hammond, 1996; Kahneman y Frederick, 2002, 2005). En segundo lugar, se ha indicado que no disponer de **juicios previos**

que permitan defender o justificar la evaluación es otro factor que permite que el procesamiento de tipo 2 opere y controle la respuesta, ya que es más probable que el nuevo juicio emitido no se encuentre sesgado por los juicios previos (Haidt, 2001, 2008). Por lo tanto, la continua experiencia y relación con el entorno permite la activación del procesamiento de tipo 1. Chaiken et al. (1989) y Kruglanski y Gigerenzer (2011) han sugerido que la activación del procesamiento heurístico también depende de si están cognitivamente disponibles los heurísticos en la memoria y si son de fácil acceso y recuperación.

Un tercer factor modulador serían las **características de la tarea**. Varios trabajos han señalado el **grado de dificultad de la tarea** como variable que afecta la activación de uno u otro proceso cognitivo (Alter et al., 2007; Hammond, 1996; Kahneman y Frederick, 2002, 2005). Lieberman (2003) indica que ante una tarea compleja o ante un caso nuevo del que no se disponga de experiencia previa para su resolución, el procesamiento de tipo 1 no puede operar y falla en su funcionamiento. En estos casos, el procesamiento de tipo 2 será activado (Evans y Over, 1996; Logan, 1988; Norman y Shallice, 1986; Toates, 2006; Schneider y Shiffrin, 1977; Shiffrin y Schneider, 1977). Por su parte, Kahneman y Frederick (2002, 2005) indican que el **formato en que se presente la información** afecta al modo en que se procesa y se resuelve la tarea. Así, los sujetos tienden a operar con el procesamiento de tipo 2 cuando (1) la información de la tarea se presenta en forma de frecuencias y no en forma de probabilidades o porcentajes, ya que es interpretado como más representativo o vívido de la realidad y (2) el estilo de la letra en que se presenta la información es más compleja de leer o interpretar (por ejemplo, letra en cursiva), ya que se exige mayor atención en su procesamiento (Alter et al., 2007). Por otro lado, Toates (2006) señala que el grado de procesamiento **emocional que la tarea** demanda es otra característica de la tarea que puede influir en la activación de los procesos. Cuando el procesamiento emocional es de bajo nivel se activa más fácilmente el procesamiento de tipo 2, ya que es emocionalmente neutro, más lógico y abstracto (Epstein, 2003; Frankish, 2010; Metcalfe y Mischel, 1999; Nisbett et al., 2001).

Finalmente, el **estrés** es otro de los factores examinados. Metcalfe y Mischel (1999) introducen esta variable como una variable situacional que modula la activación de uno u otro proceso cognitivo. Entienden el estrés como el estado de ansiedad derivado de las características ambientales a la que se ven sometidos los sujetos. Estos autores asumen que, ante altos niveles de estrés, el sistema “cálido” (procesamiento de tipo 1) domina al sistema “frío” (procesamiento de tipo 2) potenciando su funcionamiento. Ante bajos niveles de estrés, el sistema “frío” es capaz de operar e incluso de mejorar su funcionamiento. Sin embargo, este se vuelve disfuncional cuando la presencia de estrés alcanza altos niveles.

2.3.6. *Evolución Biológica*

Numerosos artículos han sugerido que el procesamiento de tipo 1 y de tipo 2 no solo se diferencian en cuanto a sus características, funciones y elementos con los que operan, sino que, además, son diferentes desde el punto de vista evolutivo. Se afirma que el procesamiento de tipo 1 opera independientemente de la consciencia y es más primitivo y básico que los procesos dependientes de la consciencia y del control de los impulsos (Reber, 1993; Kahneman y Frederick, 2002, 2005). Además, es evolutivamente más antiguo y se define como una característica innata generalmente adaptativa como resultado de la evolución natural y de la experiencia con el entorno (Evans y Over, 1996; Bargh y Chartrand, 1999; Loewenstein, 1996; Stanovich, 1999, 2011). Epstein (2003) indica que el sistema experiencial (procesamiento de tipo 1) es el resultado de millones de años de evolución y que los humanos compartimos este modo de procesamiento con los animales. Por ello, el proceso intuitivo se caracteriza porque produce un comportamiento adaptativo que no envuelve ni una cognición sofisticada ni la consciencia o intención de la conducta (Toates, 2006).

Sin embargo, se ha indicado que el desarrollo evolutivo del sistema racional es más reciente en el tiempo y es único de la especie humana (Epstein, 2003; Frankish, 2010; Haidt, 2001, 2008). Se afirma que el procesamiento de tipo 2 emerge del propio proceso basado en estímulos y que el principal objetivo de su desarrollo es ayudarnos a adaptarnos a las nuevas necesidades de la vida moderna (Epstein, 2003; Hammond, 1996; Schneider y Shiffrin, 1977; Sloman, 1996; Toates, 2006).

3. OTROS MODELOS COGNITIVOS

Varios investigadores han propuesto modelos cognitivos que, a diferencia de los MPD, proponen la existencia de un único proceso cognitivo (Kruglanski y Gigerenzer, 2011) o de una multiplicidad de ellos (Neisser, 1963). En este último caso, se puede diferenciar entre modelos que postulan tricotomías (Pollock, 1989, 1991; Stanovich, 2011), donde frecuentemente descomponen el procesamiento de tipo 2 en dos procesos o sistemas y modelos que postulan múltiples procesos (Evans, 2008; Conrey et al., 2005; Glöckner y Witteman, 2010; Stanovich, 2011) o un continuo de los procesos, donde el procesamiento intuitivo y analítico conforman los polos opuestos de ese continuo (Hammond, 1996; Reber, 1993). A continuación, se sintetizan los principales modelos.

3.1. **Modelo de Proceso Único**

Kruglanski y Gigerenzer (2011) propusieron un modelo unificado de los dos procesamiento cognitivos. El modelo es aplicado al contexto social del juicio y evaluación de mensajes y se configura bajo las siguientes premisas: (1) Los juicios intuitivos y los juicios

deliberados se basan en reglas que pueden ser la base de los dos tipos de juicio. Algunas de ellas son conscientemente aplicadas y otras de forma inconsciente. Las reglas pueden ser maleables, es decir, pueden ser aprendidas, desaprendidas, olvidadas y recuperadas de la memoria. (2) Cuatro factores intervienen en la elección de las reglas: las características de la tarea, la memoria del sujeto, la capacidad de atención y la motivación del individuo y la racionalidad ecológica de las reglas (es decir, la percepción de que la aplicación de la regla conducirá a un mejor resultado). (3) Existe una relación recíproca entre el grado de dificultad de aplicación de la regla y las capacidades del individuo. (4) La precisión de los juicios intuitivos y deliberados dependerán de la racionalidad ecológica de las reglas.

3.2. Modelos Tricotómicos

3.2.1. *Teoría General de la Racionalidad (OSCAR: General Theory of Rationality) de Pollock (1989, 1991)*

Esta teoría se configura en base a dos tipos de procesamiento diferenciados: los procesos no intelectuales (*nonintellectual processes*), definidos como procesos rápidos e inflexibles que operan a través de las asunciones del medio, y la intelección (*intellection*) que se encarga del razonamiento explícito, es más lento y exhaustivo, requiere de más recursos cognitivos y es extremadamente flexible. La intelección, a su vez, se divide en dos modos de razonamiento, racionalidad teórica (*theoretical rationality*), referida al razonamiento de las creencias, y la racionalidad práctica (*practical rationality*), concerniente a todos los aspectos que no se incluyen en el razonamiento teórico. Ambos razonamientos interaccionan entre sí. Por ejemplo, para operar, el razonamiento práctico se basa en el conocimiento de la situación que trata de procesar y ese conocimiento es facilitado por el razonamiento teórico. Además, presupone que cualquier razonamiento práctico parte del razonamiento teórico. El razonamiento práctico causa la formación de las intenciones y los deseos, mientras que el razonamiento teórico se encarga de la formación de las creencias. En su conjunto, la intelección se encarga de: (1) operar en aquellas situaciones en las que el razonamiento rápido e inflexible no puede actuar y (2) controlar y anular los resultados del razonamiento lento e inflexible cuando sea necesario.

3.2.2. *Modelo Híbrido de Procesamiento de Evans (2009)*

Evans (2009) propuso el modelo híbrido que sugiere la existencia de un procesamiento de tipo 3 asociado al sistema 3. El modelo híbrido se desarrolla con el objetivo de explicar el espacio temporal que se da en los modelos paralelos-competitivos desde que, el procesamiento de tipo 1 finaliza su proceso hasta que el procesamiento de tipo 2 se completa y uno de ellos controla la respuesta. Por lo tanto, el procesamiento de tipo 3 se define como un sistema preconsciente y es el responsable de reclutar a la memoria operativa y al procesamiento de tipo 2

y resolver los conflictos entre ambos procesos. En los modelos paralelos-competitivos, el funcionamiento del procesamiento de tipo 1 y del procesamiento de tipo 2 se inicia de forma simultánea y en paralelo. No obstante, el primero, al ser más rápido, completa su proceso antes que el segundo. En el transcurso de este espacio temporal, el sistema 3 decide si acepta la respuesta del procesamiento de tipo 1, sin ser previamente revisada por el procesamiento de tipo 2, o espera a que el sistema 2 complete su procesamiento. Este modelo se asocia con los estudios de Thompson y colaboradores (Thompson, 2009; Thompson et al., 2013) donde proponen la metacognición como un tercer tipo de procesamiento cognitivo, responsable de regular el procesamiento total de razonamiento y delegar al procesamiento de tipo 2 la responsabilidad única del procesamiento deliberado en sí mismo.

3.2.3. Modelo Tripartito de Procesamiento Cognitivo de Stanovich (2011)

Este modelo se configura sobre la noción de que el razonamiento humano se lleva a cabo a través de tres mentes: la mente autónoma, la mente algorítmica y la mente reflexiva. La primera de ellas se encarga del procesamiento de tipo 1 y las dos siguientes ejecutan el procesamiento de tipo 2. De este modo, las diferencias individuales en la racionalidad humana vienen establecidas por las diferencias individuales en ambos niveles de procesamiento analítico. El nivel algorítmico hace referencia a la habilidad cognitiva y se encarga de la ejecución de tareas mentales. Las medidas de inteligencia fluida son indicadores de las diferencias individuales en este nivel de procesamiento, siendo evaluados mediante tareas de desempeño máximo. Sin embargo, el nivel reflexivo hace referencia a las disposiciones de pensamiento (*thinking dispositions*) y se encarga de llevar a cabo tareas que implican priorización de objetivos, manejar y gestionar los valores y creencias y la autorregulación epistémica en general. Las medidas sobre disposiciones de pensamiento indican diferencias individuales en el nivel reflexivo de procesamiento y son, frecuentemente, evaluadas bajo condiciones de desempeño típico. El modelo sostiene que el concepto de racionalidad es un concepto más amplio que el de inteligencia, donde se engloban las características y propiedades de la capacidad cognitiva (mente algorítmica) y de las disposiciones de pensamiento (mente reflexiva; Stanovich 2009, 2011).

3.3. Modelos de Múltiples Procesos

3.3.1. Modelos que Definen el Procesamiento como un Continuo

Reber (1993) diferencia dos cogniciones que deben ser vistas como funcionalmente complementarias y cooperativas, denominadas proceso implícito o conocimiento tácito y proceso explícito. Reber asume que no hay una clara definición de hasta dónde llega el procesamiento implícito y donde comienza el procesamiento explícito, así que define el proceso de pensamiento

como un continuo donde ambos tipos de cognición se encuentran en los polos. Por lo tanto, los diversos tipos de procesamiento incluyen elementos de un pensamiento consciente e inconsciente.

En esta misma línea, Hammond (1996) crea su propio modelo de procesamiento aplicado al dominio de la toma de decisión y juicio. Aunque diferencia entre dos formas de procesamiento, uno intuitivo y otro analítico, asume que no existe una dicotomía en cuanto al modo de procesar, sino que el procesamiento es un continuo que se mueve entre el proceso intuitivo y el proceso analítico. El modelo se configura en base a los siguientes axiomas: (1) las diversas formas de cognición pueden ser ordenadas de acuerdo con un continuo, lo que permite tantos procesos como posibles combinaciones entre la intuición y el análisis. (2) El “sentido común” es el tipo de procesamiento habitualmente utilizado y ninguna tarea se lleva a cabo mediante un procesamiento “puro”, es técnicamente imposible que ninguna tarea dependa exclusivamente de procesos automáticos o de procesos controlados. (3) Con el tiempo, la actividad cognitiva puede moverse o desplazarse a lo largo del continuo. La cognición exitosa inhibe el movimiento, sin embargo, el fracaso lo estimula. Cuando se falla en la respuesta, la actividad cognitiva se mueve a lo largo del continuo hacia la intuición. (4) La cognición humana opera de acuerdo con el principio de coherencia (que no haya contradicciones), correspondencia (mediante asociaciones) y reconocimiento (mediante situaciones consideradas similares). Estos principios pueden sesgar los juicios y la toma de decisiones, por lo que, cualquier cognición puede aplicar heurísticos en su funcionamiento.

3.3.2. Modelos que Postulan una Multiplicidad de Procesos

3.3.2.1. Modelo de Glöckner y Witteman (2010)

Este modelo defiende la existencia de múltiples procesos de tipo 1 que han sido sintetizados en 4 tipos: (1) *Intuición asociativa basada en el simple proceso de aprendizaje y recuperación*: como resultado de la interacción con el entorno, la información puede ser recordada de forma consciente e inconsciente. Los mecanismos de aprendizaje empleados dan lugar a diferentes tipos de intuición: (a) intuiciones que son meros sentimientos de atracción y aversión; (b) intuiciones que representan estados afectivos de arousal y (c) intuiciones que reflejan la activación de previos patrones de conducta exitosos. (2) *Intuición de unión basado en la comparación con ejemplares o prototipos*: la intuición puede depender de procesos complejos de aprendizaje y recuperación de la memoria asociando situaciones y objetos con ejemplares o prototipos y recuperándolos o recordándolos mediante características de la memoria. (3) *Intuición acumulativa basada en la evidencia automática*: las opciones en la toma de decisión se seleccionan a través de procesos automáticos de integración de la información percibida y almacenada en la memoria. Esta información es constantemente inspeccionada y cuando alcanza el umbral de certeza, la opción es seleccionada. (4) *Intuición constructiva basada en las representaciones mentales*: la información se puede construir mediante representaciones mentales

basadas en la información que está activa en la memoria, a través de nuevas interpretaciones de la información y mediante la combinación de elementos de forma creativa e imaginativa.

3.3.2.2. Modelo de Conrey, Sherman, Gawronski, Hugenberg y Groom (2005)

Los autores postulan un modelo cuádruple de procesamiento que identifica 4 procesos cognitivos. Dos de ellos están asociados con el procesamiento de tipo 2: (1) procesos cognitivos conducentes a evaluar y extraer información o determinar la respuesta correcta (identificados como *discriminability*) y (2) procesos de autorregulación o autocontrol (denominados *overcoming bias*). Los otros dos están asociados con el procesamiento de tipo 1 y serían: (3) procesos encargados de la activación espontánea de las asociaciones que capturan la atención hacia la secuencia conductual activada (*association activated*) y (4) procesos encaminados a facilitar una respuesta cuando el procesamiento controlado no puede operar o falla al hacerlo (proceso denominado *guessing*).

El funcionamiento de este modelo se explica del siguiente modo: ante la presencia de un estímulo, el sujeto puede (a) activar automáticamente las asociaciones relacionadas con el estímulo o (b) no hacerlo. Si activa las asociaciones, el procesamiento analítico puede entrar en funcionamiento, controlando la activación automática de las asociaciones (proceso cognitivo de *overcoming bias*). El funcionamiento de este proceso está condicionado al grado en que el sistema analítico es capaz de activar el procedimiento controlado (proceso cognitivo *discriminability*) y este, a su vez, depende de la cantidad de atención prestada al estímulo y de la capacidad cognitiva del sujeto. Si la tendencia a responder con sesgos es superada, entonces el proceso cognitivo analítico (*discriminability*) conduce la respuesta. Por el contrario, si la tendencia a responder con sesgos no es superada, entonces la respuesta automática (proceso cognitivo *association activated*) produce la respuesta. Por último, cuando las asociaciones no son automáticamente activadas y no se dispone de una respuesta correcta, opera el proceso cognitivo de *guessing*, es decir, el sujeto supone o adivina.

3.3.2.3. Otros modelos

Otros modelos que han postulado 3 tipos de procesamientos o sistemas cognitivos también han afirmado que el procesamiento de tipo 1 engloba una multiplicidad de procesos automáticos. A este respecto, Stanovich (2009, 2011) sugiere el acrónimo TASS (*The Autonomous Set of Systems*) para describir al conjunto heterogéneo de sistemas autónomos que dan lugar al procesamiento de tipo 1 y que operan automáticamente en respuesta a la presencia de un estímulo. Asimismo, sugiere que estos procesos pueden incluir acciones de regulación de las emociones, acciones encaminadas a resolver problemas de adaptación, procesos de

aprendizaje implícito y condicional, procesos de asociación y procesos adquiridos de forma automática a través de la experiencia.

Por su parte, Evans (2009) propone que el sistema 1 está conformado por una multiplicidad de sistemas que recogen varios procesos implícitos. Los sistemas pueden ser de dos tipos: (1) sistemas autónomos que controlan el comportamiento directamente sin la necesidad de una atención controlada y (2) los sistemas preatentivos que proveen contenido a la memoria operativa y determinan qué información entra en el proceso analítico. En relación con los sistemas preatentivos, el autor desarrolló el modelo híbrido expuesto anteriormente.

4. CRÍTICAS A LOS MODELOS DUALES DE PROCESAMIENTO

Algunas de las características de los MPD han sido recientemente cuestionadas. Las críticas han señalado aspectos ya comentados en los apartados previos de este capítulo. Así, por ejemplo, una primera crítica ha sido enumerada por Evans y Stanovich (2013). Los autores indican que la **vaguedad en la definición de las etiquetas** y la **multiplicidad de denominaciones** para referirse a los tipos de procesamiento es un primer aspecto que implica controversia en la formulación y unificación de los MPD (véase apartado 2). Cada una de las denominaciones, asociativo/basado en reglas, heurístico/analítico, automático/controlado, intuitivo/racional; implican connotaciones semánticas propias. De ahí, que cada tipo de procesamiento esté asociado con propiedades muy concretas. Sin embargo, las denominaciones deberían ser neutras sin connotar características de los procesamientos ni si están conformados por uno o múltiples procesos cognitivos.

Una segunda crítica se relaciona con la particularidad de que los MPD **asumen una perfecta covarianza** o relación entre las características asociadas a los procesamientos, a pesar de que esto no tiene por qué ser así. Spunt (2015) indica que las categorías de procesamiento cognitivo asumen una perfecta correlación entre las propiedades de cada proceso. De modo que, ningún proceso puede ser categorizado como automático o controlado evaluando su posición en relación con una de las propiedades o rasgos. Si un proceso es inconsciente deberá ser asumido como no controlado, ineficiente y no intencional como consecuencia de esa perfecta relación. Recientemente, investigaciones en este sentido sugieren que las categorías son ortogonales y, por lo tanto, diferentes métodos son requeridos para testar las hipótesis relacionadas con las propiedades de los procesos. Por ejemplo, Evans y Stanovich (2013) indican que existe una vaguedad en la definición de consciencia, pues afirman que ambos tipos de procesamiento pueden recoger aspectos conscientes o no conscientes. La sugerencia de que el procesamiento de tipo 2 está basado en reglas y el de tipo 1 no, también ha sido criticada por algún modelo (Kruglanski y Gigerenzer, 2011).

La tercera crítica ha sido relacionada con la falta de **consistencia o coherencia entre los propios modelos** desarrollados. Smith y DeCoster (2000) establecieron que los MPD difieren

principalmente en dos propiedades: (1) en cuanto al factor o aspecto que enfatiza o modera el funcionamiento de uno u otro tipo de procesamiento (véase apartado 2.3.5) y (2) en cuanto al modo en que ambos tipos de procesamiento operan: simultánea, independiente o secuencialmente (véase apartado 2.3.4). Respecto al punto 1, algunos autores defienden que la motivación es un aspecto relevante para que el procesamiento de tipo 2 opere: si el sujeto no está lo suficientemente motivado el procesamiento de tipo 1 tenderá a ser empleado. Otros autores aseveran que las diferencias individuales moderan la activación del procesamiento analítico. Sin embargo, hay quien defienden que tanto aspectos motivacionales como la capacidad cognitiva afectan a la activación de los sistemas. En cuanto al modo en que ambos tipos de procesamiento interactúan, los MPD diferencian entre modelos que interactúan o que son independientes (dimensión funcional) y modelos que operan en paralelo o de forma secuencial (dimensión espacio-temporal). Para tratar de solucionar esta falta de consistencia se ha tratado de unificar los diversos MPD.

La cuarta crítica subyace del hecho de que nuevas investigaciones sugieren la existencia de un **procesamiento continuo o de múltiples procesos frente a un procesamiento dicotómico** (véase apartado 3). Existe, por lo tanto, una falta de consenso en cuanto a si el procesamiento cognitivo es esencialmente dicotómico o si cada tipo de procesamiento está a su vez conformado por múltiples procesos (Conrey et al., 2005; Evans, 2009; Glöckner y Witteman, 2010; Stanovich, 2009, 2011). Asimismo, aunque con menos éxito hasta el momento, se ha llegado a postular la posibilidad de que sea un único proceso cognitivo el que dé lugar al razonamiento o procesamiento cognitivo (Kruglanski y Gigerenzer, 2011). En este sentido, Evans y Stanovich (2013) diferencian entre los términos tipos y modos de procesamiento. Los modos de procesamiento hacen referencia a los diferentes estilos cognitivos aplicados en un procesamiento de tipo 2 y, frecuentemente, se confunden con el término tipos de procesamiento. El término tipo de procesamiento hace referencia a la terminología empleada para denominar el procesamiento dicotómico que es cualitativamente diferente. Como el procesamiento de tipo 2 es flexible, permite aplicar diversos estilos cognitivos (Evans y Stanovich, 2013). El modelo tripartito de procesamiento desarrollado por Stanovich (2011) se fundamenta en esta distinción.

Por último, otro aspecto que ha sido criticado de los MPD ha sido la **valencia atribuida a cada tipo de procesamiento**. Tradicionalmente, se ha identificado al procesamiento de tipo 1 como más negativo, al estar vinculado con la comisión de sesgos cognitivos (que, frecuentemente, resultan en respuestas erróneas o menos acertadas) y con la aplicación de prejuicios y estereotipos (Evans y Stanovich, 2013; Hammond, 1996; Lieberman, 2003). Algunos autores han afirmado que la intuición es un peligro, y no solo por la propensión a cometer sesgos sino por la excesiva confianza que los sujetos muestran en sus respuestas intuitivas (Hammond, 1996). No obstante, aunque en menor medida, algunos autores han defendido una visión más positiva del procesamiento intuitivo (Evans, 2008; Glöckner y Witteman, 2010). Así, por ejemplo, Brainerd y Reyna (2001) admiten que, aunque el procesamiento difuso (proceso de tipo 2) es superior al

procesamiento del quiz de la cuestión (proceso de tipo 1) en términos de precisión, este último aporta numerosas ventajas cognitivas cuando opera. Por ejemplo, (1) la memoria del procesamiento de tipo 1 es más persistente a lo largo del tiempo que la memoria del procesamiento de tipo 2, en cuanto al grado de detalles aportados y, por lo tanto, es más fácil acceder a la información almacenada; (2) la memoria del proceso de tipo 1 es más generalizable a diferentes formas de razonamiento que la memoria del proceso de tipo 2, orientada a un problema en particular; (3) el razonamiento de tipo 1 es menos complicado, exige menos esfuerzo, consume menos tiempo y es propenso a cometer menos errores que el razonamiento de tipo 2 al operar con detalles más precisos; (4) aunque en general, el funcionamiento del razonamiento de tipo 2 aporta una respuesta más precisa y ajustada a la realidad, el procesamiento de tipo 1 puede incrementar su potencial de precisión al cometer menos errores por operar con información menos compleja.

El procesamiento de tipo 2 ha estado vinculado a una visión del modo de razonamiento más positiva, al asociarse con una mayor deliberación propia, únicamente, del ser humano, por ser un procesamiento más lento, consciente y lógico de la información y por no estar vinculado al uso de juicios estereotipados (Evans y Stanovich, 2013). No obstante, recientes modelos han afirmado y evidenciado que el procesamiento de tipo 2 no está libre de la comisión de sesgos cognitivos, por lo que su tradicional visión de superioridad comienza a ser objeto de debate (Guthrie et al., 2007; Kahneman y Frederick, 2002, 2005; Stanovich, 1999, 2011). En cualquier caso, resulta evidente que ambos tipos de procesamiento son importantes y necesarios y, aplicar uno u otro dependerá de la situación en que nos encontremos (Hammond, 1996). Ante una situación de peligro inminente, resulta impensable emplear un procesamiento más deliberado y lento como es el proceso de tipo 2, mientras que, en situaciones en las que hay que tomar la decisión de qué coche o casa comprar, parece necesario que sea el sistema analítico el que opere. De acuerdo con Epstein (2003), “afortunadamente no hay necesidad de escoger entre los sistemas. Cada uno de ellos tiene sus ventajas y desventajas y las ventajas de uno pueden compensar las limitaciones del otro. (j) Donde sí tenemos una opción es en mejorar nuestra capacidad para usar cada uno de ellos, y usarlos de manera complementaria”.

CAPÍTULO 2.

Reflexibilidad Cognitiva (RC) y Test de Reflexibilidad Cognitiva (TRC)

1. INTRODUCCIÓN

Una de las ilustraciones más representativa del procesamiento dual es, posiblemente, el proceso de respuesta al Test de Reflexibilidad Cognitiva (*Cognitive Reflection Test*; en adelante TRC) desarrollado por Shane Frederick (2005; Kahneman y Frederick, 2002). Este test se caracteriza porque cada uno de sus ítems provoca una respuesta intuitiva que viene a nuestra mente de forma inmediata y que, si la aceptamos sin más deliberación, estaremos respondiendo de forma errónea a los ítems. El correcto desempeño en el TRC exige la activación del procesamiento deliberado de tipo 2 que inhibe la respuesta intuitiva del sistema 1 y encuentre la respuesta correcta. De este modo, el TRC introduce por primera vez el constructo de reflexibilidad cognitiva (*cognitive reflection*; en adelante RC) como la capacidad o disposición cognitiva a inhibir la respuesta inmediata de nuestra mente (Frederick, 2005).

Desde su desarrollo, se han llevado a cabo numerosos estudios con 2 finalidades principales: testar la validez del TRC y examinar el procesamiento cognitivo que subyace en el desempeño en el test.

El objetivo de este capítulo es presentar una revisión de la RC, el TRC y el proceso cognitivo necesario para completar los ítems que conforman el test. En primer lugar, se introduce el TRC y el constructo de RC. En segundo lugar, se presentan los ítems que conforman el test y las características de corrección y administración de la medida. A continuación, se sintetiza el proceso cognitivo para responder al test y el proceso metacognitivo del que son conscientes los sujetos cuando responden correcta e incorrectamente a los ítems. Seguidamente, se analizan otros determinantes de la activación del procesamiento de tipo 1 y de tipo 2 en el desempeño en el TRC. Finalmente, se sintetizan las principales ventajas y limitaciones del test.

2. CONCEPTO DE REFLEXIBILIDAD COGNITIVA (RC) Y TEST DE REFLEXIBILIDAD COGNITIVA (TRC)

Frederick (2005; Kahneman y Frederick, 2002) diseñó el TRC, un test compuesto por 3 ítems, con el objetivo de medir las diferencias individuales en la RC. El término RC fue definido por el autor como “la capacidad o disposición para resistirse a responder con la primera idea que viene a nuestra mente” (Frederick, 2005, pág. 35). El TRC se caracteriza fundamentalmente porque presenta problemas matemáticos que (1) exigen invertir tiempo y esfuerzo mental, examinando y resolviendo las cuestiones planteadas y (2) tienden a provocar una respuesta intuitiva e inmediata, que es engañosamente plausible, aunque en realidad sea incorrecta (Baron et al., 2015; Hardisty y Weber, 2009; Pennycook, Cheyne, Koehler y Fugelsang, 2015; Pinillos, Smith, Nair, Marchetto y Mun, 2011). La respuesta inicial va acompañada de un sentimiento de exactitud o certeza (*Feeling of Rightness*, FOR; Thompson, Turner y Pennycook, 2011) que la envuelve de una engañosa plausibilidad incluso cuando finalmente se conoce la respuesta normativa.

El TRC se define como una medida que evalúa diferencias individuales en impulsividad cognitiva (sistema 1) frente a un pensamiento deliberado (sistema 2; Böckenholt, 2012; Fehr y Huck, 2015; Graffeo, Polonio y Bonini, 2015; Mastrogiorgio y Petracca, 2014; Oechssler, Roeder y Schmitz, 2009; Welsh, Burns y Delfabbro, 2013). Consecuentemente, Frederick propuso un concepto general de la RC basado en el MPD de Kahneman y Frederick (2002) y se configuró como una simple, pero importante medida de procesamiento dual (Guthrie et al., 2007).

Frederick identifica 2 características relacionadas con el sistema 2 de procesamiento para responder de forma normativa a los ítems: (1) la capacidad para dirigir los resultados del sistema 1 (función de monitorización) y (2) la capacidad para inhibir el funcionamiento del sistema 1 (función de inhibición). Por lo tanto, el rasgo de RC surge para operacionalizar la idea de que el sistema 2 es capaz de cumplir con estas dos funciones, empleando el tiempo necesario para completar la evaluación del funcionamiento del sistema 1 (Campitelli y Labollita, 2010; Del Missier, Mäntylä y De Bruin, 2011; Moritz et al., 2014). Así, algunos autores han definido el TRC como una medida de disposición de pensamiento (Campitelli y Labollita, 2010; Shenhav et al., 2011), de modo que se puede diferenciar entre los sujetos que tienden a emplear el sistema 1 y, por ende, disponen de un sistema 2 “perezoso”, y los que tienden a emplear el sistema 2, activando los mecanismos de monitorización e inhibición del procesamiento de tipo 1 (Baron et al., 2015; Campitelli y Labollita, 2010).

Dentro del marco teórico de los MPD, la explicación de respuesta al TRC se basa en los MPD de tipo por defecto-intervencionista. Frederick sugirió que, cuando los sujetos tratan de resolver los ítems, el sistema 1 genera una opción rápida e intuitiva. El sistema 2 dirige el resultado del sistema 1 e inhibe y corrige su funcionamiento si fuese necesario. Toplak et al. (2011) definieron la activación del sistema 1 por defecto como un procesamiento cognitivo avaro

o miserable (*cognitive misers*) común a los sujetos que responden a este test. La miseria cognitiva es la tendencia de los sujetos a cometer errores cognitivos por contestar con la primera respuesta que recupera nuestra mente, sin deliberar en mayor profundidad. Las personas, frecuentemente, no comprueban ni piensan detenidamente en la respuesta, asumiendo respuestas erróneas como respuestas correctas y plausibles. Por lo que la capacidad cognitiva general no es garantía suficiente de precisión en la toma de decisión. Se requiere de una capacidad cognitiva específica que permita detectar que la respuesta inmediata e intuitiva es errónea y sustituirla por otra respuesta más ajustada y deliberada. Esta es la capacidad a la que se le denomina RC (Alós-Ferrer y Hügelschäfer, 2012; Narayanan y Moritz, 2015; Pennycook, Cheyne, Barr, Koehler y Fugelsang, 2014b; Primi, Donati, Chiesi y Morsanyi, 2018; West, Meserve y Stanovich, 2012). Así, se configura el TRC como la medida de miseria cognitiva por excelencia (Toplak et al., 2011, 2014).

La alternativa de respuesta intuitiva es la alternativa más fuerte o inicialmente dominante, pero simultáneamente es la respuesta errónea o no normativa, por lo que se exige su inmediata anulación (Toplak et al., 2011). De este modo, responder correctamente a los ítems en el TRC requiere la supresión y evaluación de la respuesta impulsiva. Esta supresión, podría, por lo tanto, estar relacionada con la función de inhibición directa e indirectamente evaluada en medidas de funcionamiento ejecutivo (Graffeo et al., 2015; Lesage, Navarrete y De Neys, 2013; Stanovich, 2011). Por ejemplo, Lesage et al. (2013) definen y emplean el TRC como un *proxy* de las medidas de diferencias individuales en capacidad cognitiva ejecutiva general. Gómez-Chacón, García-Madruga, Rodríguez, Vila y Elosúa (2011) definieron el TRC como una medida de la capacidad metacognitiva de los sujetos en la solución de problemas. Es decir, definen el TRC como un test que mide la capacidad de las personas para controlar su comportamiento de forma consciente y reflexiva y la capacidad para inhibir la primera respuesta que nos viene a la mente (Finucane y Gullion, 2010; Gómez-Chacón, García-Madruga, Vila, Elosúa y Rodríguez, 2014). Del mismo modo, De Missier et al. (2011) indican que la RC está asociada con la regulación cognitiva y la definen como la capacidad para procesar de modo analítico en lugar de intuitivo cuando sea necesario en la toma de decisión. Sin embargo, estos autores van un paso más allá, indicando que la RC implica, además, la capacidad de detectar y aplicar en cada momento el estilo de pensamiento más adecuado.

3. TEST DE REFLEXIBILIDAD COGNITIVA (TRC)

El TRC (Frederick, 2005; Kahneman y Frederick, 2002) se compone de tres problemas matemáticos donde los participantes deben reproducir una respuesta para cada uno de los ítems (véase Tabla 4). Frecuentemente, los ítems son definidos como problemas *quasi-matemáticos* por lo básico o la sencillez de su contenido matemático (Barr, Pennycook, Stolz y Fugelsang, 2015; Cheyne y Pennycook, 2013). Así, una vez que la respuesta inmediata viene a nuestra mente, se

puede verificar fácilmente si la respuesta, en efecto, es correcta (Meyer et al., 2015; Oechssler et al., 2009). Además, Frederick (2005) indica que los ítems del TRC son considerados fáciles, en el sentido de que la solución es sencillamente comprendida una vez que esta es explicada. Por lo tanto, el correcto desempeño en el TRC parece no solo exigir el rasgo de RC (funciones metacognitivas de monitorización e inhibición) sino también la capacidad matemática o aritmética. De forma que, si somos capaces de detectar que la respuesta automática es errónea e inhibirla, encontrar la respuesta correcta requiere de cálculos numéricos relativamente fáciles (Campitelli y Gerrans, 2014; Pennycook et al., 2014b; Barr et al., 2015). En este sentido, la terminología *quasi-matemáticos* podría enfatizar la relevancia de las funciones metacognitivas de monitorización e inhibición en el desempeño en el TRC sobre la relevancia de las capacidades matemáticas (véase capítulo teórico 3).

Tabla 4.

Test de Reflexibilidad Cognitiva (TRC)

1. Un bate y una pelota cuestan 1.10 euros. El bate cuesta un euro más que la pelota. ¿Cuánto cuesta la pelota? ----- céntimos
(respuesta normativa: 5 céntimos) (respuesta intuitiva: 10 céntimos)
2. Si 5 máquinas tardan 5 minutos en hacer 5 aparatos, ¿cuánto tiempo tardan 100 máquinas en hacer 100 aparatos? ----- minutos
(respuesta normativa: 100 minutos) (respuesta intuitiva: 5 minutos)
3. En un lago, hay una zona de hojas de nenúfar. Cada día, la zona se duplica en tamaño. Si se tardan 48 días para que la zona cubra todo el lago, ¿cuánto tiempo haría falta para que la zona de nenúfares cubra la mitad del lago? ----- días
(respuesta normativa: 47 días) (respuesta intuitiva: 24 días)

Asimismo, el TRC se establece como una medida objetiva (y no subjetiva) donde las puntuaciones alcanzadas en el test están relacionadas con el comportamiento observado (Oechssler et al., 2009). En general, se ha configurado como una medida de desempeño típico, sin tiempo límite establecido para responder, aunque algún estudio empírico ha empleado el TRC como una medida de desempeño máximo (Guillen y Hakimov, 2014; Ring, Neyse, David-Barett y Schmidt, 2016). En término medio, los sujetos necesitan entre 2.5 y 3 minutos para completar el test (Cokely, Galesic, Schulz, Ghazal y Garcia-Retamero, 2012; Oechssler et al., 2009; Primi, Morsanyi, Chiesi, Donati y Hamilton, 2015).

Aunque Frederick diseñó el TRC de respuesta abierta con el objetivo de que los sujetos reprodujesen por sí mismos la respuesta a cada ítem, diversos autores han propuesto versiones de respuesta múltiple (Böckenholt, 2012; Gómez-Chacón et al., 2011; Morsanyi, Busdraghi y Primi, 2014). Cuando el tipo de respuesta es de formato abierto, existen 3 posibles respuestas: (a) la

respuesta normativa o correcta, (b) la respuesta intuitiva errónea (es decir, la respuesta plausible pero incorrecta que nos ofrece inmediatamente nuestra mente) y (c) otras respuestas erróneas no intuitivas (por ejemplo, por errores de cálculo; Campitelli y Gerrans, 2014; Pennycook, Cheyne, Koehler et al., 2015). Estudios previos han evidenciado que los participantes tienden a producir, en su mayoría, la respuesta intuitiva errónea y la respuesta normativa, siendo la proporción de respuestas erróneas no intuitivas menor (Campitelli y Gerrans, 2014; Koriat, 2017; Liberali, Reyna, Furlan, Stein y Pardo, 2012; Primi et al., 2018). Por ejemplo, Campitelli y Gerrans (2014) encontraron que, de una muestra total de 2,019 ciudadanos australianos, en promedio respondieron .94 ítems correctamente ($SD = 1.06$). El promedio de respuestas incorrectas fue de 1.61 ($SD = 1.04$) y de otro tipo de respuestas .45 ($SD = .69$).

Por su parte, las versiones del TRC con formato de respuesta múltiple presentan opciones que tratan de reproducir las respuestas de formato abierto. Así, se suelen utilizar 3 opciones de respuesta: (a) una respuesta correcta o normativa, (b) una respuesta errónea intuitiva o heurística y (3) otra respuesta errónea no intuitiva o respuesta *imposible de determinar*. Diversos autores han indicado que diseñar una versión del TRC de elección forzosa podría minimizar el número de respuestas incorrectas debido a errores aritméticos, haciendo que se generen las respuestas correctas más fácilmente (Gómez-Chacón et al., 2011; Morsanyi et al., 2014; Primi et al., 2015). Los resultados de un experimento llevado a cabo por Bourgeois-Gironde y Van Der Henst (2009) mostraron que los sujetos obtuvieron mejores puntuaciones en el desempeño del problema del bate y la pelota cuando tuvieron que seleccionar una de las opciones de respuesta que cuando tuvieron que reproducir por sí solos la respuesta. Gangemi, Bourgeois-Gironde y Mancini (2015) encontraron resultados en esta misma línea.

Las respuestas en el TRC son, habitualmente, codificadas por los investigadores como 0, si la respuesta es incorrecta, y 1, si la respuesta es correcta y la puntuación total del test es el sumatorio del número de ítems respondidos correctamente. De modo que la puntuación final refleja las diferencias individuales en la capacidad para ignorar la respuesta impulsiva (Baldi, Iannello, Riva y Antonietti, 2013; Bergman, Ellingsen, Johannesson y Svensson, 2010; Frederick, 2005; Welsh et al., 2013). Esta estrategia de corrección no distingue entre las respuestas incorrectas, según sean intuitivas u otro tipo de errores (Pennycook, Cheyne, Koehler et al., 2015).

Recientemente, se ha indicado que el TRC también puede medir impulsividad cognitiva, entendida como el estilo cognitivo de los sujetos que alcanzan puntuaciones bajas en el TRC (Graffeo et al., 2015). En este caso, la estrategia de corrección del test consiste en sumar el número de ítems respondidos incorrectamente (Agnew y Harrison, 2017; Brosnan, Hollinworth, Antoniadou, y Lewton, 2014; Cueva et al., 2015; Piazza y Sousa, 2014; Ponti y Rodríguez-Lara, 2015; Shenhav et al., 2011). Esta estrategia no distingue entre las respuestas incorrectas de otro tipo y las respuestas correctas e invierte el uso estándar del TRC (Frederick, 2005), siendo el TRC intuitivo una medida de impulsividad cognitiva que captura la incapacidad para suprimir la

respuesta errónea (Ponti y Rodríguez-Lara, 2015). De este modo, el TRC se configura como una medida de doble utilidad que permite medir la RC y la impulsividad cognitiva (Pennycook, Cheyne, Kochler et al., 2015). Baron et al. (2015) definieron el TRC como una medida de reflexibilidad-impulsividad (RI) y estimaron su magnitud en base al sumatorio de las puntuaciones z del desempeño en el TRC y el tiempo de respuesta al mismo. Estos autores sugirieron que el índice RI puede ser más útil que la medida de precisión (o desempeño) del TRC por sí sola.

En relación con el tipo de posibles respuestas a los ítems y el modo de corrección del test, Böckenholt (2012) propuso un modelo de respuesta de miseria cognitiva (*The Cognitive-Miser Response Model*) donde indica que el procesamiento cognitivo es secuencial y no paralelo. El modelo propuso la existencia de 3 tipos de respuestas diferentes. Cuando se inicia el procesamiento cognitivo, el primer paso es llegar a una respuesta intuitiva a través del procesamiento del sistema 1. Esta respuesta será la respuesta final si los sujetos no disponen de suficiente control para inhibir esa respuesta y activar el procesamiento de tipo 2 que anule la respuesta intuitiva y busque una respuesta correcta. Sin embargo, al igual que el procesamiento de tipo 1, el procesamiento de tipo 2 puede desencadenar una respuesta no correcta si el heurístico de sustitución es inapropiadamente aplicado y de este modo, el procesamiento deliberado falla en su funcionamiento. Por lo tanto, la respuesta intuitiva puede darse por el procesamiento del sistema 1 y por el procesamiento deliberado del sistema 2. No obstante, tanto el procesamiento deliberado como el intuitivo pueden ofrecer otras respuestas erróneas diferentes de los errores intuitivos.

Sinayev y Peters (2015) llevaron a cabo dos estudios experimentales aplicando el modelo de respuesta propuesto por Böckenholt (2012). Los autores concluyeron que para resolver los ítems del TRC se requiere, inicialmente, que los participantes piensen en la respuesta intuitiva incorrecta. Seguidamente, sus respuestas individuales son determinadas en un proceso de dos pasos: uno de reflexibilidad cognitiva y otro de capacidad numérica. Es decir, si la reflexibilidad cognitiva es lo suficientemente alta, la persona comprueba la respuesta, determina si esta es errónea y procede al siguiente paso. Para responder correctamente, la persona tiene que disponer de conocimiento sobre el conjunto de ecuaciones matemáticas apropiadas a llevar a cabo, así como disponer de la capacidad para resolver las operaciones de aritmética. Si la capacidad matemática no es lo suficientemente alta, emergerá la respuesta errónea no intuitiva. Por lo mismo, los dos pasos del proceso se registran mediante los tres diferentes tipos de respuesta en el TRC: respuesta intuitiva errónea, respuesta errónea no intuitiva y respuesta correcta. En base a esta teoría, los autores asumen que la puntuación total en el TRC es la suma de respuestas no intuitivas, es decir, de respuestas correctas y erróneas no intuitivas (paso 1 = reflexibilidad cognitiva). La medida de cálculo es el número total de respuestas correctas en el TRC (paso 2 =

capacidad numérica). Se considera que el TRC es respondido correctamente cuando se alcanza el paso 2.

En resumen, el TRC es un test objetivo de desempeño típico y que está compuesto por tres sencillos ítems matemáticos que exigen básicas operaciones de aritmética para reproducir la respuesta correcta. Requiere alrededor de 3 minutos para ser completado y fácilmente se pueden verificar las respuestas. Tres tipos de respuestas son comunes en el test: respuestas correctas, respuestas erróneas intuitivas y otros errores no intuitivos. Sin embargo, frecuentemente, se utiliza el sumatorio total de respuestas correctas para indicar la capacidad o disposición para monitorizar e inhibir el procesamiento de tipo 1. A pesar de la sencillez de sus ítems, estudios previos siguen mostrando que los sujetos tienden a fallar sistemáticamente al responder a los ítems del TRC (Frederick, 2005; Pennycook, Cheyne, Kochler et al., 2015), lo que anima a seguir examinando el procedimiento cognitivo que subyace de la respuesta a los ítems.

En los apartados siguientes se sintetizan los principales resultados empíricos sobre los procesos cognitivos y metacognitivos para desempeñar con precisión el TRC. Asimismo, se presentan evidencias científicas sobre otros posibles determinantes de la activación de los modos de procesamiento y sus efectos en el desempeño en el test.

3.1. Procesamiento Cognitivo para Responder al TRC

Consideremos el primer ítem del TRC: “Un bate y una pelota cuestan 1.10 euros. El bate cuesta un euro más que la pelota. ¿Cuánto cuesta la pelota?”. En el proceso cognitivo de respuesta al ítem, la respuesta inmediata que acude a la mente es 10 céntimos. Sin embargo, esta respuesta intuitiva es errónea, como se puede demostrar aplicando una mayor deliberación. Si la pelota cuesta 10 céntimos, como nuestro sistema 1 nos sugiere, y el bate cuesta un euro más entonces el bate tendría que costar 1.10 euros. Si sumamos los precios de ambos objetos, el bate y la pelota en su conjunto costarían 1.20 euros, no 1.10 euros como indica el enunciado. Por lo tanto, la respuesta normativa sería 5 céntimos, el bate costaría 1.05 euros y, conjuntamente, ambos objetos valdrían 1.10 euros (Guthrie et al., 2007).

De los diversos ítems que componen el TRC (Frederick, 2005), el ítem del bate y la pelota ha sido el más ampliamente considerado e investigado en la literatura científica, no solo porque este ítem ha sido debatido con anterioridad al desarrollo del TRC (Kahneman y Frederick, 2002, 2005) sino también, porque sistemáticamente un porcentaje significativamente mayor de sujetos fallan al responder el ítem (Brañas-Garza et al., 2015; Fehr y Huck, 2015; Skagerlund, Lind, Strömbäck, Tinghög y Västfjäll, 2018; Toplak, West y Stanovich, 2017; Weinhardt, Hendijani, Harman, Steel y Gonzalez, 2015). La explicación de por qué tanta gente falla al responder el ítem, se ilustra en base a la noción del heurístico de sustitución de atributos (De Neys et al., 2013; Kahneman y Frederick, 2002, 2005; Morewedge y Kahneman, 2010; Szollosi et al., 2017). La sustitución de atributos se da cuando la gente, ante un problema complejo, responde de forma

inmediata o intuitiva con una respuesta fácil. Es decir, tratamos de minimizar la carga cognitiva que exige la resolución de una pregunta compleja sustituyéndola intuitivamente por una pregunta más fácil (misericia cognitiva). La sustitución de atributos se produce cuando se evalúa el atributo objetivo que es relativamente inaccesible para la mente (disponibilidad) mediante la asignación de un atributo heurístico relativamente accesible y asociado con el atributo objetivo (representatividad; Kahneman y Frederick, 2005). Así, por ejemplo, en el ítem del bate y la pelota, el atributo de sustitución consiste en reemplazar la relación crítica “más que” por una afirmación absoluta que es más simple. De modo que la afirmación “el bate cuesta 1.00 euro más que la pelota” es sustituido por la afirmación absoluta “el bate cuesta 1.00 euro”, motivo por el que los sujetos tienden a responder 10 céntimos en lugar de 5 (De Neys et al., 2013; Szollosi et al., 2017).

Macchi y Bagassi (2012) afirman que la responsabilidad del fracaso en el desempeño de este ítem radica en la estructura retórica (modo de organizar la información de un texto) del enunciado del problema. La formulación del ítem parece indicar que el precio del bate es conocido (1.00 euro) al crear la ilusión de que el propio enunciado ofrece la respuesta al problema planteado. De este modo, la frase “el bate cuesta 1.00 euro más que” se interpreta en el sentido de “el bate cuesta 1.00 euro” ignorando la información “más que”. La teoría de sustitución de atributos demuestra que los errores intuitivos en el TRC reflejan un proceso de monitorización vago (Kahneman y Frederick, 2005; Szollosi et al., 2017). Por ello, Macchi y Bagassi (2012) afirman que las diferencias individuales en el desempeño en el TRC son debidas a un tipo muy concreto de RC: la competencia pragmática; ya que la tarea se interpreta a la luz del contexto de presentación de la información en el enunciado y no se abstrae o se descontextualiza de ella.

Liberali et al. (2012) sostienen que, aunque el fracaso en el desempeño en el TRC es debido a la redacción del enunciado, también podría ser debido a pensar superficialmente en el problema, sin detenerse ni profundizar en su significado. Los autores denominaron a los errores intuitivos del TRC errores por unión de palabras (*verbatim matching errors*) o errores sin sentido o mecánicos (*mindless errors*) y defienden que su ocurrencia viene provocada por el empleo de las propias palabras del enunciado para responder a los ítems, sin comprender realmente su significado. Así, Liberali et al. (2012) argumentan que la respuesta intuitiva del ítem de las máquinas¹ se proporciona en vista de que el enunciado indica que 5 máquinas tardan 5 minutos en hacer 5 aparatos. Por lo tanto, los sujetos inmediatamente asumen que 100 máquinas tardarán 100 minutos en hacer 100 aparatos. No obstante, los sujetos con puntuaciones altas en el TRC son capaces de suprimir y modificar la respuesta intuitiva errónea al comprobar que esta no puede ser correcta. Por lo tanto, una característica crítica del TRC es que los sujetos que responden de forma

¹ Si 5 máquinas tardan 5 minutos en hacer 5 aparatos, ¿cuánto tiempo tardan 100 máquinas en hacer 100 aparatos?

intuitiva no han comprobado previamente sus respuestas (Frederick, 2005; Kahneman y Frederick, 2005; Liberali et al., 2012).

Consecuentemente, diversos estudios han comprobado si modificaciones o manipulaciones en la redacción del ítem del bate y la pelota determinan la activación de un modo u otro de procesamiento cognitivo. Bourgeois-Gironde y Van Der Henst (2009) llevaron a cabo un estudio con dos condiciones experimentales. En la condición control, los participantes cubrieron el ítem del bate y la pelota desarrollado por Frederick (2005) con la única variante de que los elementos “bate y pelota” fueron sustituidos por “sello y sobre”. En la condición experimental, los participantes cubrieron el mismo ítem, pero se les indicó expresamente que debían tener en cuenta dos cuestiones para responder: (1) en conjunto los dos ítems valen 1.10 euros y (2) la diferencia en el precio entre ambos ítems es de 1 euro. Los resultados mostraron que no hubo diferencias significativas entre ambas condiciones en el porcentaje de personas que respondieron correcta e incorrectamente el ítem. Por lo que la manipulación de la redacción del enunciado para conminar a los participantes a prestar más atención a las reglas o premisas del problema no surtieron efecto en la mejora del desempeño en el problema.

Macchi y Bagassi (2012) también examinaron esta cuestión. En la condición control, los estudiantes cubrieron el problema del bate y la pelota desarrollado por Frederick (2005) y en la condición experimental, los estudiantes cubrieron el ítem: “un bate y una pelota cuestan 1.10 euros. El bate cuesta un euro más que la pelota. Encuentra el precio del bate y el de la pelota”. Los resultados mostraron que, en la condición experimental, el 90% de los sujetos respondieron correctamente el ítem, mientras que en la condición control, lo hicieron solamente el 10%. Estas diferencias fueron estadísticamente significativas demostrando que pedir el precio de ambos elementos facilita la correcta resolución del problema.

Del mismo modo, Mata, Schubert y Ferreira (2014) encontraron que los sujetos que fueron capaces de detectar algún cambio crítico en la redacción de los ítems del TRC tras haber respondido a los mismos mostraron una mayor probabilidad a responder correctamente los problemas que los sujetos que no detectaron cambios.

En su conjunto, estos estudios sugieren que hacer manipulaciones o modificaciones en la redacción del enunciado de los ítems del TRC puede ayudar a la activación del procesamiento tipo 2 y, por consiguiente, determinar el desempeño en los ítems. Asimismo, sugieren que leer con mayor detenimiento los ítems puede ayudar a responder de forma correcta.

No obstante, estudios previos también han mostrado que, en relación con los problemas análogos de los ítems del TRC como es el caso del problema del plátano y el pastel² (Frederick, 2005), los sujetos tienden a responder correctamente en mayor proporción el ítem análogo que el

² *Un plátano y un pastel cuestan 37 céntimos. El plátano cuesta 13 céntimos más que el pastel. ¿Cuánto cuesta el pastel?*

problema original (Bourgeois-Gironde y Van Der Henst, 2009; Frederick, 2005; Mastrogiorgio y Petracca, 2014; Thompson et al., 2011). Sin embargo, al tratarse de problemas equivalentes donde la retórica del enunciado es idéntica parece lógico pensar que el modo de redacción del enunciado no sea el principal determinante de las diferencias en el desempeño de estos ítems. Liberali y colaboradores (2012) indican que las diferencias en el desempeño en el ítem del bate y la pelota y su ítem análogo del plátano y el pastel pueden ser debidas a que el enunciado del segundo no proporciona una respuesta inmediata, un valor obvio que permita ser atribuido como respuesta al ítem (como si ocurre en el primero) y, por ello, este problema invita a la activación del sistema 2 (Bourgeois-Gironde y Van Der Henst, 2009). Los autores hipotetizan que, si el problema indicase que el plátano cuesta 30 céntimos, posiblemente los participantes responderían que el pastel cuesta 7 céntimos. Mastrogiorgio y Petracca (2014) propusieron que el determinante del desempeño de ambos ítems podría ser la fisionomía de los números envueltos en el enunciado de los problemas. Cuando ambos ítems fueron administrados experimentalmente en previos estudios, no se prestó atención a las magnitudes envueltas en ambos problemas. El problema del bate y la pelota incluye cantidades de 1.10 y 1.00 euros, mientras que el problema análogo envuelve cantidades de 0.37 y 0.13 euros, lo que indica que no comparten estrictamente una homogeneidad aritmética (diferentes magnitudes numéricas), aunque sí una homogeneidad algebraica (mismos cálculos matemáticos). Emplear números enteros más complejos (o menos prominentes) que no invitan a hacer cálculos matemáticos mentales con suficiente grado de confianza podría definir el éxito o fracaso en la respuesta a los ítems (Frederick, 2005).

En su estudio, Mastrogiorgio y Petracca (2014) suministraron dos versiones diferentes del problema del bate y la pelota siendo, desde el punto de vista matemático isomórficos, es decir, homogéneos desde una visión algebraica (ambos problemas implican el mismo procedimiento o cálculo matemático para llegar a la solución correcta) y aritmético (ambos problemas envuelven magnitudes numéricas similares o que sus diferencias son consideradas insignificantes: $1.10 - 1.00$ y $1.10 - 1.01$ dólares para el problema original del bate y la pelota y su versión isomórfica, respectivamente). Asimismo, en el diseño de la nueva versión se empleó exactamente el mismo enunciado para controlar posibles efectos de la redacción del enunciado en la respuesta a la pregunta. Los resultados del estudio mostraron diferencias significativas en el porcentaje de respuestas correctas. El 56.7% de los participantes respondieron correctamente a la pregunta del bate y la pelota original (Frederick, 2005), mientras que el 80% de los sujetos lo hicieron en la versión isomórfica. De este modo, se mostró que, controlando los resultados por la estructura algebraica y aritmética del problema y el enmarque o enunciado de la tarea, los números o sus magnitudes no son los responsables de la activación del sistema 2, sino que los numerales (los símbolos numéricos o la fisionomía del número) son los determinantes del desempeño en los problemas. El problema del bate y la pelota envuelve la misma estructura algebraica que el problema del plátano y el pastel. Sin embargo, las magnitudes del enunciado son completamente

diferentes, lo que puede responder a porqué más gente falla al responder a la primera y no tanto a la segunda pregunta.

Mastrogiorgio y Petracca (2014) señalaron que, desde un punto de vista estrictamente matemático, dos operaciones matemáticas que difieren en las magnitudes (aunque compartan la misma estructura algebraica) implican una carga o esfuerzo de cálculo diferente, pero si los dos problemas envuelven magnitudes equivalentes además de una estructura algebraica idéntica el esfuerzo de cálculo es el mismo. Sin embargo, desde el punto de vista de la cognición humana, dos problemas que comparten la misma estructura lógica y las mismas magnitudes podrían ser percibidas como tareas que requieren un esfuerzo de cálculo diferente, dependiendo de los ejemplos específicos que envuelva la tarea. De acuerdo con la teoría prominente (Albers, 2002), las operaciones aritméticas que envuelven números prominentes o enteros (10, 20, 15) requieren menos esfuerzo cognitivo o menos esfuerzo de cálculo que las mismas operaciones de cálculo que impliquen números no prominentes (como es el caso del problema del plátano y el pastel). De este modo, los números prominentes son facilitadores del cálculo aritmético en el sistema decimal, provocando el automatismo aritmético. El número 1 utilizado en el problema del bate y la pelota es más prominente que el número 1.01 usado en la versión modificada del problema, a pesar de la insignificante diferencia de la magnitud de ambos números (Mastrogiorgio y Petracca, 2014). De esta forma, los símbolos numéricos de los problemas del TRC se podrían configurar como posibles determinantes del desempeño en el TRC.

3.2. Metacognición del Procedimiento de Respuesta a los Ítems

Con el objetivo de profundizar en el conocimiento procedimental de respuesta a los ítems del TRC, se han llevado a cabo diversos estudios sobre la metacognición de respuesta al test. La metacognición es la cognición sobre la cognición, es decir, es la capacidad de los seres humanos para analizar la consciencia del proceso cognitivo. Los principales estudios han confirmado que las diferencias en la metacognición de respuesta al TRC son debidas a las diferencias en los modos de pensamiento (Mata, Ferreira y Sherman, 2013; Mata y Almeida, 2014).

La asimetría metacognitiva ha sido demostrada por diversos estudios. Por ejemplo, los resultados del estudio de Mata, Ferreira et al. (2013) reflejaron que los participantes que respondieron de forma intuitiva al TRC no indicaron haber pensado en otra respuesta previamente. Sin embargo, los participantes que respondieron correctamente indicaron, en su mayoría, haber pensado previamente en la respuesta intuitiva. Mastrogiorgio y Petracca (2014) pidieron a los participantes de su estudio que explicasen el proceso de razonamiento del problema del bate y la pelota una vez completado el ítem. Los sujetos que respondieron correctamente el problema afirmaron haber pensado inicialmente en la respuesta intuitiva y, al comprobar si la respuesta era correcta, la modificaron por la respuesta reflexiva. Del mismo modo, Mata y Almeida (2014) encontraron que los sujetos deliberados, frecuentemente, indican que han

considerado la respuesta intuitiva antes de llegar a la respuesta deliberada y sugieren que la respuesta deliberada puede venir a la mente solo después de haber reconsiderado la respuesta inicial.

Travers, Rolison y Feeney (2016) examinaron el tiempo y proceso de respuesta a los ítems del TRC. Los autores encontraron que cuando los participantes seleccionaron la respuesta intuitiva, reflejaron menos evidencias de conflicto durante el procesamiento cognitivo. El cursor del ordenador se desplazó más veces cuando los participantes escogieron la respuesta correcta que cuando seleccionaron la respuesta impulsiva. Además, el cursor se movió más rápidamente hacia la respuesta intuitiva que hacia la respuesta correcta. La distancia mínima entre el cursor y la respuesta heurística, cuando seleccionaron la respuesta correcta, fue significativamente menor que la distancia entre el cursor y la respuesta correcta, cuando seleccionaron la respuesta heurística. Por otro lado, cuando los participantes proporcionaron la respuesta correcta, invirtieron más tiempo en la región de la respuesta heurística que en la región de cualquier otra opción de respuesta. Esto sugiere que se consideró la respuesta heurística antes de alcanzar la respuesta correcta. Por lo mismo, la inhibición de la respuesta intuitiva es un importante predictor del buen desempeño en el TRC. Sin embargo, los participantes que dieron la respuesta heurística no presentaron mayor probabilidad de poner el cursor en la región de la respuesta correcta que en cualquier otra opción de respuesta, lo que sugiere que los sujetos intuitivos no prestan la suficiente atención a la respuesta deliberada.

Szaszi et al. (2017) encontraron que, en su mayoría, los sujetos que responden correctamente a los ítems en el TRC inician su deliberación con la respuesta correcta y no tanto con la respuesta intuitiva. Asimismo, el tiempo final invertido en responder al TRC fue significativamente mayor en el grupo de sujetos que comenzaron la deliberación con la respuesta incorrecta, presumiblemente porque este grupo necesitó llevar a cabo operaciones mentales extra para llegar a la respuesta correcta. Del total de sujetos que respondieron incorrectamente a los ítems el 39% llevaron a cabo algún tipo de comportamiento reflexivo tras indicar su respuesta inmediata (por ejemplo, al mostrar dudas sobre la respuesta), mientras que el 61% aceptaron la respuesta intuitiva sin mayor deliberación. Además, los sujetos del grupo reflexivo que tuvieron dudas al indicar la respuesta intuitiva emplearon significativamente más tiempo en resolver los problemas que los sujetos del grupo no reflexivo. Los autores concluyeron que la tendencia individual a usar mecanismos heurísticos en el procesamiento cognitivo podría no ser la única causa responsables de fallos en el procesamiento o pensamiento humano, sino que un problema de *mindware* por la falta de conocimiento declarativo o de reglas estratégicas necesarias para resolver los problemas también podría explicar los errores en el TRC y en el pensamiento humano en general (Sinayev y Peters, 2015; Szaszi et al., 2017). Muchos sujetos no son capaces de encontrar la solución correcta a los ítems del TRC incluso deliberando acerca de su primera respuesta (Stanovich, Toplak y West, 2008; Toplak et al., 2011).

Por lo tanto, la asimetría metacognitiva propone que los sujetos deliberados son más conscientes de las respuestas intuitivas y deliberadas que los sujetos intuitivos, ya que muchos de ellos responden correctamente después de haber considerado la respuesta intuitiva; mientras que los impulsivos solo son conscientes de la respuesta intuitiva (Mata y Almeida, 2014). Trouche et al. (2014) encontraron que el 45% de los participantes que respondieron de forma intuitiva al ítem del bate y la pelota cambiaron su respuesta al conocer la explicación de respuesta de otro participante que resolvió correctamente el ítem. Mientras que ningún participante que contestó correctamente modificó su respuesta al leer la argumentación de un participante que respondió intuitivamente. Esto parece indicar que los sujetos intuitivos no son conscientes de la respuesta deliberada del ítem. En este mismo sentido, el estudio llevado a cabo por Krawczyk y Sylwestrzak (2018) mostró que el TRC correlacionó con el número de veces que los sujetos cambiaron la opción de respuesta, sugiriendo que los sujetos con puntuaciones más altas en RC se mostraron más dispuestos a cambiar sus decisiones, aparentemente porque son más conscientes del proceso metacognitivo y son más capaces de reconocer que tomaron decisiones inferiores. Los resultados defienden, una vez más, la ventaja metacognitiva que los sujetos deliberados presentan sobre los sujetos intuitivos.

No obstante, el hecho de que los sujetos intuitivos no sean tan conscientes de la respuesta deliberada no implica necesariamente que ellos no detecten el proceso de sustitución de atributos en la respuesta a los ítems. Con el objetivo de comprobar esta hipótesis, De Neys et al. (2013) diseñaron una nueva versión del problema del bate y la pelota que no exigiese el proceso de sustitución (denominado ítem de no conflicto o versión control). Este ítem se caracteriza por no incorporar en su redacción la frase “más que”³. Posteriormente, los participantes indicaron el grado de seguridad con las respuestas facilitadas. Los resultados mostraron diferencias significativas en la respuesta a ambas versiones del ítem. El 21% de los participantes respondieron correctamente el ítem original frente a un 98% que lo hicieron en el ítem de no conflicto. A este mismo resultado llegaron Johnson, Tubau y De Neys (2014) a pesar de que, en su estudio experimental, manipularon la carga cognitiva de los participantes. Del mismo modo, De Neys et al. (2013) hallaron que el nivel de seguridad de los participantes que respondieron de forma incorrecta el ítem original fue significativamente más bajo que el nivel de confianza presentado al responder el ítem de no conflicto. Esto sugiere que los participantes fueron sensibles al proceso de sustitución y no completamente inconsciente de la naturaleza de la respuesta incorrecta.

Con el objetivo de complementar la investigación de De Neys y colaboradores (2013), Szollosi et al. (2017) llevaron a cabo un estudio experimental en el que los participantes cubrieron dos ítems del TRC, uno en versión original y otro en versión conflicto (extraído de De Neys et al., 2013). Los resultados mostraron que, para la condición congruente (TRC de no conflicto),

³ Una revista y un plátano cuestan 2.90 dólares. La revista cuesta 2 dólares. ¿Cuánto cuesta el plátano?

tanto los que indicaron que verificaron su respuesta como los que indicaron que no, el grado de confianza en la respuesta fue muy alto (75.8% vs. 82.2%, respectivamente). En la condición incongruente (TRC original), el nivel de confianza de los sujetos que indicaron que no verificaron la respuesta fue menor (38.5%) que los que indicaron que sí verificaron la respuesta (59.5%). A pesar de haber comprobado la respuesta, los sujetos de la condición incongruente se sintieron menos seguros de sus respuestas que los de la condición congruente. De modo que se encontró un efecto de la condición sobre el grado de seguridad en la respuesta. Se obtuvieron resultados similares en la consciencia sobre la dificultad de verificar las respuestas dadas. A los sujetos que cubrieron los ítems de no conflicto les resultó más fácil de verificar la respuesta correcta que a los sujetos que cubrieron los ítems en la versión original. Además, se encontró que cuando los sujetos indicaron que los problemas fueron más difíciles de verificar estuvieron menos seguros de sus respuestas. Sin embargo, Mata, Ferreira et al. (2013) no encontraron diferencias significativas en la estimación de la dificultad de realizar el TRC entre sujetos deliberados e intuitivos.

El experimento 2 de Szollosi et al. (2017) mostró que, tanto en la condición congruente como en la condición incongruente, los sujetos que respondieron correctamente a los ítems indicaron, mayoritariamente, que verificaron sus respuestas. Sin embargo, los sujetos que respondieron intuitivamente en la condición incongruente reportaron, en su mayoría, no haber verificado sus respuestas o haber fallado en su comprobación. Este resultado apoya la idea de que una de las posibles causas del mal desempeño en el TRC sea un problema de *mindware* o falta de conocimiento para resolver los problemas. Asimismo, encontraron que, en la condición congruente, los sujetos que verificaron o no verificaron sus respuestas se mostraron mucho más seguros de su exactitud que aquellos que fallaron en la verificación. Sin embargo, en la condición de incongruencia, los sujetos que verificaron sus respuestas se mostraron mucho más seguros (84.2%) que los sujetos que no verificaron sus respuestas (54.1%) o que fallaron en su verificación (15.8%). La verificación correlacionó con el grado de seguridad en la respuesta, mostrándose más seguros aquellos que consiguieron comprobar la respuesta. Los resultados mostraron que los sujetos encontraron los ítems de la versión incongruente más difíciles de verificar que los de la versión congruente. La correlación entre la dificultad de comprobar la respuesta del ítem y el grado de seguridad en la respuesta fue significativa y negativa.

Por su parte, Koriati (2017) encontró que los participantes que respondieron los ítems de no conflicto lo hicieron predominantemente de forma correcta y se mostraron más seguros de la exactitud de sus respuestas en esos ítems que en los ítems de conflicto. Asimismo, indicaron que los ítems de no conflicto fueron menos engañosos que los originales y se mostraron más seguros al indicar que los ítems de no conflicto fueron engañosos que al indicar que lo fueron los ítems de conflicto.

Por lo tanto, en su conjunto, estos resultados reflejan que los participantes más reflexivos disponen de una ventaja metacognitiva sobre los participantes intuitivos, pues tienden a ser más conscientes del procesamiento cognitivo llevado a cabo en la respuesta a los ítems del TRC. Los estudios han indicado que, aunque no todos los sujetos reflexivos inician su proceso cognitivo con la respuesta intuitiva, sino que algunos parten desde un comienzo con la respuesta deliberada, estos siguen siendo conscientes de la existencia de la respuesta intuitiva inmediata, pues muestran una mayor tendencia a examinar la exactitud de sus respuestas. Asimismo, los resultados han reflejado que los sujetos intuitivos, aunque responden de forma intuitiva, no son totalmente inconscientes de los errores en sus respuestas, pues algunos tienden a mostrar síntomas de dudas y de reflexión tras responder de forma intuitiva a los ítems del TRC. Además, se muestran menos seguros de la precisión de sus respuestas en comparación con otros problemas que no implican conflicto y en, general, no suelen examinar la exactitud de las respuestas. Sin embargo, cuando lo hacen tienden a fallar en la comprobación, lo sugiere un problema de *mindware*, más que de falta de RC.

3.2.1. Experiencia Metacognitiva del Sentimiento de Respuesta Correcta y de la Percepción del Desempeño de uno Mismo en el TRC

La respuesta intuitiva, frecuentemente, va acompañada de un sentimiento o percepción de certeza y veracidad (Thompson et al., 2011). De modo que es habitual encontrarse estudios que demuestren que los sujetos que responden de forma intuitiva a los ítems del TRC se muestran más seguros de sus respuestas que los sujetos que respondieron correctamente. De acuerdo con la teoría de razonamiento metacognitiva (Thompson et al., 2011), el sentimiento de exactitud o de veracidad (*Feeling of Rightness*; FOR) es una respuesta afectiva que complementa la respuesta rápida y sesgada del sistema 1 (Gangemi et al., 2015) y, por lo tanto, puede ser un indicativo de la necesidad de un análisis adicional en nuestro procesamiento. En otras palabras, puede ser un indicativo de la necesidad de activar el procesamiento de tipo 2 para inhibir el procesamiento de tipo 1 (Thompson et al., 2011). Thompson y colaboradores (2011) hallaron una relación negativa entre el sentimiento de exactitud o veracidad (FOR) y las medidas de pensamiento analítico. Los sujetos con menor sentimiento de exactitud estuvieron asociados con una mayor inversión de tiempo en la fase de reconsideración de la respuesta inicial. Además, mostraron una mayor probabilidad a cambiar la respuesta inicial. Los juicios de sentimiento de certeza o exactitud fueron consistentemente predichos por la fluidez con la que la respuesta inicial fue producida. Por ello, evidenciaron que la respuesta intuitiva frecuentemente va acompañada de un sentido o percepción de exactitud. En sus estudios, Thompson et al. (2013) encontraron que la fluidez de la respuesta (*answer fluency*), entendida como la velocidad con que la respuesta intuitiva inicial es producida, fue un predictor fiable del sentimiento de precisión en el desempeño de la tarea. Además, los sujetos que mostraron un mayor sentimiento de precisión en las respuestas mostraron

una menor tendencia a cambiar sus respuestas tras comprobarlas y a invertir tiempo en la comprobación.

No obstante, Trouche et al. (2014) encontraron resultados opuestos. Los datos reflejaron que los participantes que dieron la respuesta correcta se sintieron más seguros que los que facilitaron la respuesta intuitiva. Del mismo modo, los sujetos que dieron la respuesta intuitiva se sintieron más seguros de sus respuestas que los que facilitaron otro tipo de respuesta errónea. Trouche et al. (2014) también encontraron que el resultado en el ítem de las máquinas fue mejor cuando la respuesta al ítem se consensuó en grupo.

Por su parte, Gangemi et al. (2015) llevaron a cabo 2 estudios donde, en lugar de emplear el indicador FOR (sentimiento de precisión o exactitud en las respuestas), emplearon el indicador FOE (*Feeling of Error*) que examina el sentimiento de error. Los resultados mostraron que el sentimiento de error fue significativamente más alto entre los participantes que completaron el ítem original cuando lo resolvieron incorrectamente que cuando lo hicieron en la versión sin conflicto y lo resolvieron correctamente. Del mismo modo, FOE fue significativamente más bajo cuando los participantes resolvieron correctamente el problema de conflicto que cuando lo respondieron incorrectamente. En el estudio 2, además, encontraron que el formato de respuesta al test puede afectar al sentimiento de error en la respuesta al TRC. Concretamente, encontraron que los sujetos que respondieron correctamente el ítem en formato de elección forzosa mostraron puntuaciones más altas en FOE que los que respondieron correctamente en un formato de respuesta abierta.

La experiencia metacognitiva del sentimiento de exactitud en las respuestas al TRC también se puede testar mediante la estimación subjetiva del propio desempeño de los participantes. A este respecto, Mata, Ferreira et al. (2013) encontraron que los sujetos que respondieron menos ítems correctamente (sujetos intuitivos) estimaron un desempeño en el TRC por encima de su desempeño real (sobrestimación de su desempeño), mientras que los sujetos que respondieron más ítems del TRC correctamente (sujetos deliberados) fueron más precisos en la estimación de su desempeño real. Ring et al. (2016) examinaron esta misma cuestión separando a los participantes por razón de sexo. Los datos reflejaron que los hombres estimaron su desempeño más alto de lo que lo hicieron las mujeres. Sin embargo, ambos grupos sobrestimaron su desempeño de forma significativa. No obstante, no se encontraron diferencias en la sobrestimación de las puntuaciones. Finalmente, Duttie (2015) encontró que los sujetos más altos en RC tienden a percibir de forma más ajustada la precisión absoluta de su desempeño en un test de capacidad cognitiva. Sin embargo, la puntuación que alcanzaron en el TRC no correlacionó con la estimación que llevaron a cabo sobre el desempeño de los demás en relación con el suyo (medida de *overplacement*).

3.2.2. *Experiencia Metacognitiva de la Percepción del Desempeño de Otros en el TRC*

Frederick (2005) encontró que cuando se les pide a los participantes que estimen la proporción de sujetos que podrían acertar los ítems, los que han respondido correctamente tienden a señalar una proporción menor que aquellos que han fallado los ítems. Por lo tanto, presentan una estimación más precisa del desempeño de otros participantes en el test. Varios estudios han replicado estos resultados. Por ejemplo, Guthrie et al. (2007) pidieron a los participantes que juzgasen si otros participantes responderían a los ítems del TRC correctamente. Los datos mostraron que aquellos sujetos que respondieron las preguntas erróneamente mostraron mayor tendencia a considerar que los demás acertarían los ítems. Fehr y Huck (2015) encontraron que los sujetos que respondieron el ítem del bate y la pelota erróneamente estimaron que el 73% de los participantes acertarían el ítem, mientras que aquellos que respondieron correctamente estimaron un porcentaje menor (52%, $p < .01$), lo que indica que los sujetos con puntuaciones más bajas en el TRC creyeron erróneamente que la respuesta correcta del ítem era muy simple. Mata, Ferreira et al. (2013) mostraron que cuando se les pidió a los participantes que estimasen la proporción de ítems que los demás participantes responderían correctamente, los sujetos deliberados estimaron una proporción más baja que los sujetos intuitivos. Además, indicaron de forma más ajustada el porcentaje de participantes que desempeñaron todos los ítems correctamente. No obstante, ambos grupos sobrestimaron el desempeño de los demás participantes, aunque los sujetos deliberados hicieron una estimación más precisa. Asimismo, se mostraron más seguros de sus estimaciones que los sujetos intuitivos ya que, al conocer ambas alternativas de respuesta, son más conscientes de la respuesta intuitiva que otros podrían ofrecer. Sin embargo, cuando se les brindó la oportunidad de modificar sus respuestas anteriores, los participantes que respondieron intuitivamente la primera y segunda vez estuvieron más seguros de la capacidad de los demás para resolver correctamente el ítem. Los participantes que mejoraron su desempeño en la segunda vez (pasaron de intuitivos a deliberados), se mostraron menos seguros de la estimación realizada sobre el desempeño de otros participantes en la segunda vez y su evaluación fue similar a la de los participantes que respondieron correctamente desde el primer momento.

En un estudio llevado a cabo por Mata, Ferreira et al. (2013), los participantes tuvieron la oportunidad de modificar sus respuestas tras observar las respuestas de otros participantes. Cuando se les mostró la respuesta intuitiva, los sujetos que respondieron de forma deliberada no cambiaron sus estimaciones de cómo los demás participantes desempeñarían el TRC. Asimismo, tampoco modificaron sus estimaciones cuando se les mostró la respuesta deliberada. Cuando a los sujetos que respondieron de forma intuitiva se les mostró la respuesta intuitiva tampoco modificaron sus estimaciones y se mostraron más seguros con las respuestas iniciales. Sin embargo, cuando se les presentó la respuesta deliberada modificaron sus estimaciones sobre el

desempeño de los demás. La diferencia absoluta de las estimaciones hechas en el momento 1 y 2 fueron mayores para los sujetos intuitivos que para los deliberados.

Por último, Ring et al. (2016) examinaron la misma cuestión, pero separando a los participantes por razón de sexo. Los datos no mostraron diferencias significativas entre la estimación que los hombres y las mujeres hicieron a cerca del desempeño de los demás participantes. Ambos grupos estimaron que su desempeño sería significativamente más alto que el de los demás participantes del estudio, lo que indica una sobre confianza en su desempeño (*overplacement*). Aunque, los hombres tendieron a sobreestimar en mayor medida su desempeño en comparación con el desempeño de los demás que las mujeres (mayor *overplacement* para los hombres, $z = -3.737, p < .001$). Del mismo modo, hombres y mujeres estimaron que su desempeño sería mejor que el de los demás hombres y solo los hombres estimaron su desempeño más bajo que el de las demás mujeres.

No hay razón alguna para creer que los sujetos deliberados tienen una habilidad social especial para inferir el desempeño de los demás y que los intuitivos carecen de ella (Mata y Almeida, 2014), por lo que la posible explicación a los resultados encontrados subyace de la ventaja metacognitiva que los sujetos deliberados presentan frente a los sujetos intuitivos. Los sujetos reflexivos son más conscientes de la respuesta intuitiva inmediata que otros pueden dar, del sentimiento de seguridad y precisión que implica y del grado real de dificultad de los ítems, lo que permite hacer una estimación más precisa del desempeño de los demás. Sin embargo, los sujetos intuitivos desconocen la alternativa de respuesta deliberada y, por ende, esperan que los demás sujetos respondan en la misma línea en que lo han hecho ellos. Por tanto, es de esperar que los sujetos intuitivos estimen que los demás responderán como ellos lo han hecho, es decir, en base a la aplicación del proceso cognitivo seguido por ellos mismos (Mata y Almeida, 2014). Estos resultados empíricos sugieren la posibilidad de que los sujetos deliberados puedan anticiparse al desempeño de los demás a través de esta ventaja metacognitiva y resultar en una ventaja estratégicas para los mismos.

En este sentido, Mendonça (2012) llevó a cabo un estudio con dos condiciones experimentales: condición yo-otros (donde los participantes, primero, completaron el TRC y seguidamente, respondieron al TRC de nuevo desde el punto de vista de cómo lo harían otros) y la condición otros-yo, (donde los participantes cubrieron el TRC desde el punto de vista de cómo lo responderían otros participantes y a continuación, lo completaron por ellos mismos). Los resultados mostraron que los sujetos respondieron a más ítems correctamente cuando respondieron por ellos mismos que cuando lo hicieron estimando las respuestas de otros. Asimismo, se encontró que los sujetos cometieron menos errores cuando cubrieron el test por ellos mismos en la condición otros-yo que cuando lo cubrieron en la condición yo-otros. Estos resultados, por tanto, sugieren que pensar en el procesamiento cognitivo seguido por otros sujetos

en el desempeño en el TRC ayuda a mejorar el desempeño propio en el test, lo que implica, a su vez, una ventaja estratégica.

Por su parte, Mata y Almeida (2014) encontraron que los sujetos que responden de forma deliberada al TRC aprovechan la información metacognitiva que otros sujetos facilitan para inferir que tipo de procesamiento cognitivo están aplicando esas personas y determinar si responderán correctamente al TRC. De esta manera obtienen una ventaja estratégica sobre los demás. Los resultados de los 3 estudios han sugerido que los sujetos que responden de forma deliberada a los ítems en el TRC muestran una estimación más precisa sobre el razonamiento de otros sujetos que los sujetos intuitivos. Bajo circunstancias que no favorecen la deliberación, los sujetos analíticos suelen ser conscientes de las alternativas de respuesta intuitivas cuando se les pide que estimen el desempeño de otras personas en el TRC. Los sujetos deliberados son, generalmente, capaces de inferir si otros pueden responder correctamente a los ítems teniendo en cuenta la información sobre el estilo de procesamiento de las otras personas. También son conscientes que responder de forma deliberada al TRC implica una mayor estimación media del tiempo invertido en responder. Por lo tanto, son más capaces de inferir si otras personas procesan de forma deliberada o intuitiva en función del tiempo invertido en el proceso.

En síntesis, los estudios experimentales de la asimetría metacognitiva en la respuesta al TRC demuestran que los sujetos que piensan de forma más deliberada (puntuaciones más altas en el TRC) tiene una ventaja metacognitiva sobre los sujetos que piensan de forma más intuitiva (puntuaciones más bajas en el TRC). La evidencia científica reflejó que los sujetos deliberados fueron más precisos que los intuitivos en la evaluación absoluta que ellos hicieron sobre su propio desempeño y sobre el desempeño de otros. Asimismo, también fueron más precisos en la evaluación comparativa de su desempeño frente al desempeño de otros. Además, la estimación de la precisión de sus respuestas a cada ítem del test fue más ajustadas o realistas que las expresadas por los sujetos intuitivos. Estos resultados mostraron el rol que el estilo cognitivo (intuitivo vs. deliberado) tiene en la consciencia metacognitiva. Por otro lado, se expuso que aquellos sujetos que cambian su estilo cognitivo de intuitivo a deliberado en el TRC también modifican sus impresiones sobre cuan competentes son los demás en el razonamiento de los problemas que presenta el test. Finalmente, los estudios sugieren que los sujetos deliberados e intuitivos no disponen de la misma consciencia metacognitiva, ya que, los primeros son conscientes desde un principio de las alternativas de respuesta intuitiva y deliberada, mientras que los segundos solo son conscientes de la respuesta inmediata que facilitan en un principio. Además, cuando a los sujetos intuitivos se les muestra la alternativa de respuesta deliberada modifican sus estimaciones sobre el desempeño de los participantes. Sin embargo, los sujetos deliberados no modifican sus estimaciones cuando se les muestra la alternativa de respuesta intuitiva.

3.3. Tiempo de Reacción o Fluidez de la Respuesta

Estudios previos han examinado la fluidez de respuesta al TRC con el objetivo de comprobar si el tiempo de respuesta invertido en los ítems podría ser un indicativo del modo de procesamiento cognitivo empleado en el desempeño en el test. Los principales resultados han señalado que es más probable que las respuestas rápidas reflejen resultados heurísticos o basados en un procesamiento de tipo 1 que las respuestas lentas (Thompson et al., 2011; Travers et al., 2016). Así, cuando se fuerza a los sujetos a responder rápidamente, existe una mayor probabilidad a responder en base a la conclusión inicial o con la respuesta inmediata de nuestra mente que cuando se permite más tiempo para reconsiderar o deliberar la respuesta (Kinnunen y Windmann, 2013; Mata y Almeida, 2014).

Kinnunen y Windmann (2013) llevaron a cabo un estudio en el que dividieron a los participantes en dos condiciones experimentales: la condición del sistema 1, donde se animó a los participantes a responder al TRC lo más rápido posible y con la primera respuesta de nuestra mente, y la condición del sistema 2, donde se les pidió a los participantes que se tomaran todo el tiempo necesario para deliberar y razonar sus respuestas. Los resultados indicaron que los sujetos a los que se les animó a deliberar sus respuestas obtuvieron mejores resultados en el test.

Asimismo, Anderson et al. (2008) compararon el desempeño y el tiempo invertido en la respuesta al TRC de un grupo de profesionales de nivel medio-alto y un grupo de trabajadores rurales. Los resultados mostraron que el porcentaje de sujetos que acertaron el ítem del bate y la pelota en el grupo de profesionales fue más elevado que el porcentaje de sujetos que acertaron en el grupo de trabajadores rurales. Además, en todos los casos, los trabajadores cualificados invirtieron, en promedio, al menos un minuto más de tiempo en responder que los trabajadores rurales. La media de respuesta de los trabajadores profesionales fue más homogénea o menos variable, lo que sugiere una tendencia de los profesionales a tomar un tiempo extraordinario para responder la pregunta. Los resultados evidenciaron que las respuestas intuitivas o emocionales requieren una inversión de tiempo menor que las respuestas cognitivas más razonadas. No obstante, el tiempo invertido en las posteriores preguntas decrece para ambos grupos de sujetos debido al efecto aprendizaje.

Böckenholt (2012) encontró que el tiempo de respuesta invertido en cada ítem del TRC fue diferente entre sí sin que las diferencias en función del sexo o del momento del día (acrofase) en que se completó el TRC fuesen significativas. Las respuestas impulsivas y erróneas fueron significativamente más rápidas que las respuestas correctas. Sin embargo, las diferencias en cuanto al tiempo invertido en dar una respuesta correcta y una respuesta errónea diferente de la impulsiva no fueron significativas, sugiriendo que durante el procesamiento de tipo 2 también se pueden cometer errores.

Por su parte, Alós-Ferrer, Garagnani y Hügelschäfer (2016) midieron el tiempo invertido en responder al TRC en sus correspondientes versiones de conflicto (ítem original) y no conflicto

(De Neys et al., 2013; donde no se requiere aplicar el heurístico de sustitución de atributos porque el enunciado del problema no recoge “más que”). Los resultados reflejaron que los participantes respondieron a las preguntas de conflicto significativamente más despacio que a las preguntas de no conflicto. Otros estudios han obtenidos resultados similares. Por ejemplo, Mata y Almeida (2014) encontraron que responder al TRC de conflicto de forma deliberada implica una mayor estimación media del tiempo invertido que responder al TRC de no-conflicto. En la integración de 4 estudios llevado a cabo por Meyer et al. (2015) hallaron que los sujetos que respondieron al TRC en la condición no fluida invirtieron significativamente más tiempo en deliberar la respuesta. No obstante, Koriat (2017) no encontró diferencias significativas en el tiempo de reacción a los problemas de conflicto y no conflicto.

3.4. Otros Determinantes del Desempeño en el TRC

La prevalente explicación de por qué una proporción tan alta de sujetos fallan al responder a los ítems del TRC se centra en la explicación de la sustitución de atributos. La sustitución de atributos requiere de la disponibilidad o accesibilidad cognitiva con la que los sujetos acceden al heurístico. En este sentido, Kahneman y Frederick (2005) definieron la accesibilidad como la facilidad (o esfuerzo) con que un particular contenido de la memoria viene a la mente. La accesibilidad está determinada conjuntamente por los mecanismos cognitivos de los sujetos, por las características de los estímulos y eventos que se evocan, así como por las características de la situación o escena (Kahneman y Frederick, 2002, 2005). De modo que la manipulación de ciertos factores, tales como características de la tarea, pueden ayudar (entorpecer) a los sujetos a aplicar el heurístico de sustitución de atributos y activar el sistema 1 (sistema 2) de procesamiento.

En este apartado se identifican los principales aspectos del entorno o las condiciones experimentales que favorecen la activación del sistema 2 y la inhibición del sistema 1 en la realización de los problemas del TRC. Es decir, se examinan las principales estrategias de “desesgo” del TRC. En primer lugar, se presentan los determinantes relacionados con las características de las tareas y a continuación se examinan los determinantes relacionados con otras características del entorno.

3.4.1. Relacionados con la Tarea

3.4.1.1. Redacción de las Instrucciones

Relacionado con la redacción del enunciado de los problemas está la formulación de las instrucciones dadas en las tareas. Bigoni y Dragone (2011, 2012) llevaron a cabo un estudio con el objetivo de examinar cómo el modo de presentar las instrucciones de las tareas de un experimento afecta al desempeño en el mismo. Los resultados mostraron que, en promedio, los alumnos comprendieron mejor las instrucciones cortas que las largas y las presentadas en formato

papel que las presentadas en una pantalla de ordenador. Asimismo, los resultados mostraron que la comprensión de las instrucciones correlacionó significativa y negativamente con la impulsividad, medido por el TRC mediante el sumatorio de las respuestas erróneas. De modo que los sujetos con más errores en el TRC comprendieron peor las instrucciones del experimento. Además, los resultados también mostraron que los sujetos impulsivos comprendieron peor las instrucciones cuando estas fueron redactadas de forma breve, salvo que se les forzara expresamente a prestar mayor atención en la lectura. De este modo, los sujetos con bajas puntuaciones en el TRC demostraron comprender mejor las instrucciones que son más extensas.

No obstante, el estudio de Fernbach, Sloman, Louis y Shube (2013) parece no apoyar los resultados anteriores. Estos autores hallaron que los sujetos con puntuaciones bajas en RC alcanzaron una menor comprensión de la explicación a cerca de las propiedades de unos productos cuando se proporcionaron más detalles, mientras que los sujetos con altas puntuaciones en RC obtuvieron una mejor comprensión de la explicación cuando fueron detalladas en mayor profundidad. Sin embargo, no se encontró una relación lineal entre la RC y el nivel de detalles, como presentaron los sujetos con puntuaciones bajas en el TRC. A niveles intermedios y altos de detalle en las explicaciones, los sujetos con altas puntuaciones en RC reflejan niveles similares de comprensión de la información. Los sujetos con altas puntuaciones en el TRC disponen de un umbral en el que el grado de comprensión no se incrementa, aunque se incremente el nivel de detalle de las explicaciones y no hay indicios de que al superar ese umbral de detalle el nivel de comprensión se degrade, como sí ocurre en los sujetos con baja RC. Asimismo, Fernbach et al. (2013) también encontraron que, cuando los participantes tuvieron que indicar su preferencia de consumo por uno de los productos, los sujetos con bajas puntuaciones en el TRC prefirieron el producto premium cuando este fue explicado con un nivel de detalle trivial mientras que los sujetos con altas puntuaciones en el TRC prefirieron el producto premium cuando este fue explicado con mayor nivel de detalle. Las preferencias por un producto u otro correlacionaron significativamente con la comprensión de la información facilitada sobre el mismo y, por lo tanto, fue determinada por la interacción entre el nivel de detalle en la explicación y las puntuaciones en el TRC.

3.4.1.2. Fluidez de la Fuente de los Ítems o Fluidez Perceptual

La fluidez perceptiva (*perceptual fluency*) es la facilidad con que los problemas pueden ser leídos y habitualmente se relaciona con el tipo de fuente utilizada en la redacción de las preguntas de los ítems (Alter et al., 2007). Diversos estudios han sugerido que llevar a cabo manipulaciones en el estilo de fuente empleado afecta a la activación de los modos de procesamiento cognitivo. Alter et al. (2007) llevaron a cabo un estudio con una muestra de 40 estudiantes universitarios. Todos los participantes completaron el TRC en una de las dos condiciones experimentales según el formato de fuente: fácil de leer o estándar (*fluent condition*)

o difícil de leer (*disfluent condition*). Los resultados mostraron que los sujetos que cubrieron el TRC en un estilo de letra más complejo acertaron más ítems que los participantes de la condición fluida. Mientras que el 90% de los sujetos de la condición fluida respondieron al menos 1 ítem de forma incorrecta, solo el 35% lo hicieron en la condición de no fluidez. Los resultados parecen sugerir que cuanto más difícil de leer sea el enunciado mayor es la tendencia a activar el sistema 2 de procesamiento y a responder más ítems correctamente, forzando al sujeto a procesar la información de forma más cautelosa.

Thompson et al. (2013) trataron de replicar los resultados del estudio de Alter et al. (2007) y encontraron que presentar el TRC en un formato de letra difícil de leer (usar letra cursiva, símbolos, etc.) no incrementa significativamente el tiempo invertido en el procesamiento. La inversión de tiempo adicional no produce cambios en las respuestas, no mejora el desempeño final ni incrementa la probabilidad de obtener más respuestas correctas. Los autores sugieren que la no fluidez perceptual alerta a los sujetos de que algo no es correcto e invierten más tiempo en su deliberación, pero muchos, simplemente, tratan de justificar su respuesta más que cambiarla. Los autores también comprobaron la posibilidad de que los resultados de manipular la fluidez perceptual variasen en función de las capacidades cognitivas. Esta hipótesis fue demostrada únicamente para los participantes con capacidades cognitivas altas. Así, los sujetos más sofisticados mostraron mejores resultados en el TRC en la condición de no fluidez perceptual que en la condición de fluidez perceptual que los sujetos con bajas puntuaciones en capacidad cognitiva.

Meyer et al. (2015) llevaron a cabo un meta-análisis de los resultados encontrados por Alter et al. (2007) y Thompson et al. (2013). Además, integraron 13 nuevos estudios desarrollados por los propios autores. Los resultados demostraron que no existen diferencias significativas en el desempeño en el TRC entre los sujetos que lo cubrieron en la fuente estándar y aquellos que lo cubrieron en una fuente no fluida. De los 17 estudios integrados, solo el estudio de Alter et al. (2007) muestra un efecto significativo de la fuente no fluida en el desempeño en el TRC y otros 5 encontraron diferencias, aunque no llegaron a ser estadísticamente significativas. Asimismo, examinaron posibles moderadores del efecto de la fuente sobre el desempeño en el TRC: (1) La capacidad cognitiva o IQ: los resultados mostraron que esta variable no produce un efecto moderador en la relación entre el desempeño en el TRC y la fluidez perceptual. (2) Posibles pistas para activar el procesamiento de tipo 2 en la realización del TRC: no se obtuvo diferencias en el desempeño entre los sujetos que cubrieron el TRC en una fuente normal de los sujetos que lo cubrieron en una fuente no fluida y que hayan recibido pistas (ayuda) antes de cubrir el test. (3) Formato de presentación del TRC (papel u ordenador): no se encontraron diferencias significativas entre cubrir el TRC en una fuente fluida y no fluida y hacerlo en formato papel u ordenador. (4) Sitio de realización del experimento: no se encontraron diferencias significativas entre cubrir el TRC en una fuente fluida y no fluida siendo el estudio de laboratorio, público u

online. (5) Conocer con anterioridad algún ítem del TRC: tampoco se encontraron diferencias significativas en el desempeño en el TRC cubierto en una fuente fluida o no fluida y el hecho de que los sujetos estuviesen previamente expuestos al test. Por último, el meta-análisis demostró que los sujetos que respondieron en la condición no fluida invirtieron significativamente más tiempo en la respuesta que los sujetos que respondieron en la condición fluida. Además, se mostró que solo los sujetos más inteligentes invirtieron más tiempo en el desempeño en el TRC en la condición de la fuente no fluida. Los autores sugieren que esta diferencia puede deberse a que: (1) los sujetos más inteligentes invierten más tiempo en la lectura del ítem o (2) otros procesos de pensamiento cognitivo más profundos son activados por la fuente no fluida, incluyendo la reflexibilidad de porqué la fuente del ítem no es fluida.

Posteriormente a este meta-análisis, Yilmaz y Saribay (2016) replicaron el paradigma de la no fluidez cognitiva. Los resultados confirmaron los resultados del meta-análisis de Meyer et al. (2015). No se encontraron diferencias significativas en el desempeño en el TRC en ambas condiciones de fluidez perceptual.

3.4.1.3. Dificultad del Idioma

Costa, Foucart, Arnon et al. (2014) examinaron si cubrir el TRC en un idioma extranjero puede afectar a los resultados alcanzados en el test. Una parte de la muestra cubrieron el TRC en su idioma nativo y la otra en el idioma extranjero. En ambos casos, no se encontraron diferencias significativas en el desempeño en el TRC para los sujetos que cubrieron el TRC en el idioma nativo o extranjero. Los resultados reflejaron que realizar los problemas en una lengua extranjera no previene la intervención espontánea del Sistema 1. Sin embargo, un estudio posterior halló que los sujetos que respondieron al TRC en un idioma extranjero (del que poseen un nivel fluido) obtuvieron mejores resultados que aquellos que lo hicieron en su idioma nativo (Costa, Foucart, Hayakawa et al., 2014), por lo que es necesaria más literatura en este sentido.

3.4.1.4. Secuencia/Orden de los Ítems

Brañas-Garza et al. (2015) llevaron a cabo una integración de 118 estudios sobre el TRC, desarrollados durante los años 2007 a 2015, en el que se acumularon una muestra total de 44,558 sujetos. El meta-análisis examinó la distribución de las puntuaciones del TRC y su relación con otras variables como, en este caso, la secuencia u orden de los ítems. Los resultados sugirieron que el orden de presentación de los ítems puede determinar el desempeño en el test. Así, los sujetos obtuvieron mejores resultados cuando la secuencia de ítems se presentó en el orden original (bate, máquinas y nenúfares) que cuando se mostró en cualquiera de las otras posibles combinaciones, siendo más probable que no respondiesen ningún ítem correctamente cuando los ítems no estuvieron formulados en el orden estándar.

3.4.2. Relacionados con el Contexto

3.4.2.1. Posición del Test Dentro del Estudio Experimental

Nuestro cerebro se alimenta del nivel de glucosa en sangre para llevar a cabo tareas cognitivas (Kuhn, Kuhn y Villeval, 2014). Cuando el esfuerzo cognitivo realizado es importante, los niveles de glucosa en sangre tienden a disminuir, con el consiguiente cansancio o fatiga. De este modo, cabe presuponer que el desempeño en las pruebas que conforman un estudio experimental puede verse afectado por los niveles de glucosa en sangre, ya que el nivel de cansancio no es el mismo al inicio, en el medio o hacia el final del estudio.

El meta-análisis de Brañas-Garza et al. (2015) examinó si el desempeño en el TRC se ve afectado por la posición del test dentro del estudio experimental. Los resultados no mostraron diferencias significativas en el desempeño en el test según su posición en el experimento. Solo se encontraron diferencias significativas entre cubrir el ítem de los nenúfares en el medio o al final del experimento. Los participantes fallaron el ítem en mayor proporción cuando este se situó al final del experimento. Sin embargo, cuando se eliminaron los estudios no experimentales del análisis, los resultados no solo mostraron diferencias significativas en el ítem de los nenúfares sino también en el ítem del bate y la pelota. Completar el TRC al final del estudio experimental puede afectar de forma negativa a sus resultados.

3.4.2.2. Formato de Presentación del Test (Lápiz y Papel vs. Ordenador)

Asimismo, el meta-análisis de Brañas-Garza et al. (2015) examinó si el formato de presentación del TRC (lápiz y papel u ordenador) determina de algún modo el desempeño en el test. Los resultados reflejaron que el 12% de la muestra respondió correctamente a los problemas cuando se administraron en papel y lápiz, mientras que el 88% respondieron correctamente cuando la tarea se realizó a través del ordenador. Asimismo, se mostró que, de los 3 ítems, únicamente el problema del bate y la pelota no se vio afectado por el formato de presentación de los ítems. El meta-análisis sugiere que implementar el TRC en el ordenador puede mejorar el desempeño en el test.

3.4.2.3. Tarea Incentivada

Los efectos de los incentivos en el comportamiento humano han sido debatidos en la literatura psicológica a lo largo de los años. En los estudios experimentales se utilizan los incentivos para conseguir maximizar el desempeño de los participantes (Brañas-Garza et al. 2015). Así, diversas investigaciones sobre el TRC han pagado a los participantes en función del correcto desempeño en los ítems (Baghestanian y Frey, 2016; Fehr y Huck, 2015; Guillen y Hakimov, 2014; Kocher, Lucks y Schindler, 2016; Noori, 2016; Ring et al., 2016). No obstante, apenas se ha examinado si el hecho de incentivar a los sujetos provoca una mejora en el resultado.

Cueva et al. (2015) sugirieron que, posiblemente, las diferencias en función del sexo en el TRC podrían ser reducidas incentivando el TRC, puesto que en los estudios donde se ha premiado monetariamente el desempeño, las mujeres han obtenido puntuaciones medias más altas que cuando no se ha hecho. Sin embargo, esta diferencia no ha sido tan pronunciada en el caso de los hombres. No obstante, a nuestro conocimiento, esta cuestión continúa estando pendiente de examinar.

Posteriormente, el meta-análisis de Brañas-Garza et al. (2015) examinó si incentivar el desempeño en el TRC afecta a la resolución de los problemas. Los resultados del análisis de regresión reflejaron que pagar a los sujetos por las respuestas correctas a los ítems no incrementa significativamente sus resultados en el test.

3.4.2.4. Feedback

Viator, Bagley, Barnes y Harp (2014) han sugerido que las intervenciones de feedback suministran a los sujetos la oportunidad de reflexionar, considerar y evaluar los atajos cognitivos y las reglas que previamente han venido utilizando. De modo que un mínimo de RC es exigido para que las intervenciones de feedback sean efectivas. Los sujetos con bajas capacidades en RC simplemente no iniciarán el procesamiento algorítmico que permita anular la respuesta autónoma, mientras que será más probable que los sujetos con mayor RC utilicen la información del feedback y consideren la posibilidad de anular los heurísticos y reglas utilizados para mejorar su desempeño. Los autores encontraron en su estudio que los sujetos con puntuaciones más altas en RC mejoraron su desempeño tras recibir feedback general sobre las respuestas correctas, mientras que los sujetos con puntuaciones bajas en RC no mejoraron su desempeño. Asimismo, Viator et al. (2014) señalaron que facilitar feedback sobre el resultado correcto al finalizar cada tarea de forma individualizada puede facilitar el aprendizaje de aquellos participantes con relativamente menor puntuación en RC.

3.4.2.5. Momento del Día (Acrofase)

Böckenholt (2012) examinó si los sujetos mostraron una mayor tendencia a inhibir sus respuestas intuitivas en función del momento del día en el que el TRC fue administrado o en función del cansancio de los participantes. De este modo, la muestra se dividió en dos grupos “sujetos de mañana”, que respondieron al TRC entre las 7 y 10 de la mañana, y “sujetos de tarde”, que respondieron al TRC entre las 7 y las 10 de la noche. En promedio, los que contestaron por la mañana completaron más ítems correctamente que los que contestaron el test por la tarde. De este modo, los datos sugieren que responder al TRC en un momento u otro del día afecta a su desempeño.

3.4.2.6. Carga Cognitiva

Algunos estudios han encontrado que manipular la carga cognitiva de los participantes durante la realización del TRC puede influir negativamente en el desempeño en el test.

Johnson et al. (2014) llevaron a cabo un estudio con tres condiciones experimentales: (1) carga cognitiva alta: donde los participantes memorizaron la posición de 4 puntos en una matriz (tarea secundaria de la memoria del punto); (2) carga cognitiva baja: donde los participantes memorizaron la posición de tres puntos en la misma matriz, colocados más próximos entre sí en un patrón más fácil de recordar; y (3) no carga cognitiva. Los datos mostraron que los participantes obtuvieron mejores resultados cuando no estuvieron sometidos a ninguna carga cognitiva. Seguidamente, los participantes obtuvieron mejores resultados en la condición de baja carga cognitiva y, finalmente, obtuvieron un peor desempeño en la condición de mayor carga cognitiva. El efecto de la carga cognitiva en la precisión en el TRC fue estadísticamente significativo.

Morsanyi et al. (2014) llevaron a cabo un estudio siguiendo el mismo diseño experimental que Johnson et al. (2014). Los resultados mostraron un efecto significativo de la carga cognitiva en el desempeño en el TRC. Los sujetos obtuvieron mejores resultados en el TRC cuando estuvieron sometidos a una carga cognitiva baja. No obstante, la carga cognitiva no tuvo un efecto directo sobre el tiempo invertido en responder a los ítems, de modo que los sujetos tardaron de forma similar en responder al test bajo condiciones de carga y no carga cognitiva. En la condición de no carga cognitiva no hubo una relación significativa entre el TRC y el tiempo de respuesta. Sin embargo, bajo la condición de carga cognitiva, el desempeño en el TRC correlacionó con el tiempo de respuesta al test.

Por su parte, Grossman, van der Weele y Andrijevik (2014) no encontraron diferencias significativas en el desempeño en el TRC en función de la carga cognitiva expuesta durante la realización del test. Concretamente, los resultados mostraron que, en promedio, los sujetos acertaron la misma proporción de ítems en ambas condiciones de carga cognitiva (alta vs. baja). Además, los autores no encontraron diferencias significativas en el tiempo invertido en realizar el test en función de la carga cognitiva.

3.4.2.7. Agotamiento del Autocontrol

El autocontrol “es la capacidad individual para anular o inhibir la tendencia a un comportamiento no deseado, como los impulsos, y frenar las actuaciones derivadas de ellos” (Tangney, Baumeister y Boone, 2004, pág. 275). El autocontrol puede estar directamente relacionado con los MPD, pues una reducción del autocontrol puede ser interpretada como un incremento del rol del sistema 1 (impulsividad) en la toma de decisión.

Kocher et al. (2016) encontraron que mermar el autocontrol de los sujetos sometiéndolos a una tarea de *stroop* no afectó a la puntuación alcanzada en el TRC. No se encontraron diferencias

significativas en el desempeño entre los participantes que fueron sometidos a un agotamiento de su control cognitivo y los sujetos que no fueron sometidos a un agotamiento del autocontrol. Este resultado se mantuvo cuando se incluyeron las observaciones del segundo experimento. Sin embargo, este último, además, sugiere un efecto indirecto del agotamiento del autocontrol en el desempeño en el TRC, ya que los resultados reflejaron que las puntuaciones en el test explicaron los beneficios obtenidos por los comerciantes, pero solo cuando su autocontrol no se vio mermado previamente.

3.4.2.8. Estrés

Simonovic et al. (2017) llevaron a cabo un estudio experimental con el objetivo de testar si el estrés puede ser un determinante del desempeño en el TRC. Durante la investigación, los participantes fueron asignados a una de las dos condiciones según la manipulación del estrés. Los participantes de la condición experimental llevaron a cabo la tarea y tuvieron que explicar su experiencia con el experimento mientras eran grabados con una cámara. Los del grupo control no estuvieron sometidos a ese estrés y desarrollaron la tarea sin la presencia de cámaras. Para medir la respuesta al estrés se tomaron diversas medidas de la presión sistólica de la sangre durante el desempeño de la tarea y durante los tiempos de descanso. Los datos mostraron diferencias significativas en las puntuaciones en el TRC entre los dos grupos experimentales, siendo los participantes de la condición experimental los que obtuvieron un peor desempeño en el TRC. Se mostró, por tanto, que el pensamiento reflexivo fue inhibido por el estrés. Sin embargo, no se encontró una correlación significativa entre el TRC y la percepción del estrés, ni entre el TRC y la presión sistólica de la sangre.

3.4.2.9. Efecto Priming Previo

Entre las estrategias de “desesgo”, la más estudiada hasta el momento ha sido si exponer a los participantes a estímulos intuitivos o reflexivos influye en la posterior puntuación alcanzada en el TRC. Bourgeois-Gironde y Van Der Henst (2009) llevaron a cabo un estudio con 120 estudiantes que cubrieron el ítem del bate y la pelota. Todos ellos fueron divididos entre la condición control y la condición experimental. En la condición experimental, los participantes completaron el ítem precedido del problema del plátano y el pastel (*efecto priming*). Los resultados mostraron que la manipulación experimental para tratar de mejorar el desempeño en el ítem no tuvo efectos sobre el resultado en el ítem. La proporción de sujetos que respondieron correctamente al problema del bate y la pelota en ambas condiciones fue el mismo, aunque en la condición experimental un porcentaje mayor de sujetos dieron una respuesta errónea diferente de la intuitiva.

En este mismo sentido, Deppe, Gonzalez, Neiman, Jacobs y Pahlke (2015) diseñaron un estudio experimental con dos condiciones. En la condición intuitiva, previamente a cubrir el TRC,

los participantes tuvieron que componer una frase con palabras relacionadas con la intuición. En la condición control las palabras empleadas en la construcción de la oración fueron neutras. Los resultados mostraron que no hubo diferencias significativas en las puntuaciones del test entre los participantes de la condición intuitiva y de la condición control. De modo que no hubo un *efecto priming* en el desempeño en el TRC. Los estudios 2, 3 y 4 obtuvieron resultados en esta misma línea.

3.5. Limitaciones del TRC

Como se ha podido observar a lo largo de este capítulo, diversos autores han señalado que el TRC es un instrumento con ciertas ventajas. Así, por ejemplo, se ha indicado la brevedad del test como una de las características más relevantes. El TRC se identifica como un test corto que puede ser administrado en 1 o 2 minutos (Frederick, 2005). Además, es fácil de completar y posee una alta validez aparente (Oechssler et al., 2009; Szaszi et al., 2017). Diversos estudios han mostrado que es el mejor o segundo mejor predictor de una amplia variedad de variables relacionadas con la toma de decisión. Por lo que su validez predictiva aparente es igual o superior a la de otros test compuestos por más ítems y que requieren de tiempo extra para su realización (Frederick, 2005; Oechssler et al., 2009; Szaszi et al., 2017).

Sin embargo, recientemente, se ha indicado que el TRC no está libre de limitaciones y que, paradójicamente, muchas de ellas son compartidas con sus ventajas. Las principales limitaciones se relacionan con sus propiedades psicométricas y parten, en gran medida, del número tan reducido de ítems del que se compone. Originalmente, los ítems del TRC se han definido como problemas matemáticos muy sencillos de resolver (Frederick, 2005), sin embargo, estudios posteriores han reflejado la **dificultad del test**. Los principales resultados han evidenciado que, en promedio, una importante proporción de sujetos resuelven uno o ningún ítem correctamente, mientras que la proporción de sujetos que resuelven 2 o más ítems correctamente es inferior (Brañas-Garza et al., 2015; Frederick, 2005; Weller et al., 2013). Frederick (2005) administró el test a 3,428 sujetos y encontró que solo el 17% de la muestra respondió a los 3 ítems correctamente, frente a un 33% que no respondió ningún ítem favorablemente. Brañas-Garza et al. (2015) integraron 118 estudios sobre el TRC y encontraron que el 38% de la muestra acumulada ($N = 44,558$) fallaron los tres ítems frente al 18% que acertaron todos los ítems. Estudios posteriores también reflejaron resultados en esta línea (Kiss, Rodríguez-Lara, Rosa-García, 2016; Skagerlund et al., 2018; Szaszi et al., 2017; van Prooijen, 2017). Por lo tanto, la mayor parte de estudios que han examinado la distribución de las puntuaciones en el TRC han mostrado valores bajos de las medias (\bar{X}) y valores de las desviaciones típicas (SD), frecuentemente, superiores o iguales a las medias alcanzadas. El Apéndice A (Tabla A1) recopila la distribución de las puntuaciones (\bar{X} , SD y tamaño de la muestra) de diversos estudios empíricos, de forma que se puede comprobar el patrón de distribución de las puntuaciones. Con el objetivo

de conocer la distribución de las puntuaciones de la población total, se ha procedido a integrar cuantitativamente las puntuaciones promedio y sus desviaciones típicas⁴.

Como se puede observar en la Tabla A2 del Anexo A, con una muestra acumulada de 66,572 sujetos, la puntuación promedio estimada para la población ha sido 1.22 ($S^2 = .17$). Este resultado prácticamente no difiere del resultado hallado en el meta-análisis de Brañas-Garza y colaboradores ($\bar{X} = 1.20$, $N = 44,558$). Asimismo, la desviación típica de las puntuaciones estimada para la población ha sido de 1.15.

Por otro lado, muy pocos estudios han reportado otro tipo de información sobre la distribución de las puntuaciones del TRC diferente de la proporción de aciertos o de las medias y desviación típica de las puntuaciones. Weller et al. (2013) reportaron que, en promedio, la muestra de 1,970 participantes acertó .83 ítems del TRC ($SD = 0.99$) pero, además, informó del valor de la mediana y la moda, así como la asimetría de las puntuaciones. Encontraron que el valor de la mediana y la moda para su muestra fue igual a cero y la asimetría de .88. Juanchich et al. (2016) encontraron una asimetría de las puntuaciones de .67 en una muestra de 384 participantes. Finalmente, Del Missier et al. (2011) obtuvieron una asimetría de 1.19 y una curtosis de 0.32 en una muestra de 212 sujetos.

Todos estos datos, en su conjunto, indican una **distribución de las puntuaciones altamente sesgada** (asimetría y curtosis positivas), es decir, las puntuaciones no se distribuyen normalmente entre la población, ya que la gran mayoría de la muestra se agrupa en el rango de puntuaciones bajas, mientras que una proporción menor se agrupa en el rango de puntuaciones medias y altas. Por lo tanto, su sensibilidad para detectar de forma precisa el nivel de capacidad del sujeto en el constructo latente podría verse limitado por el “**efecto suelo**” (“*floor effect*”) que produce el propio test. En este sentido, Weller et al. (2013) señalan la **capacidad del TRC para evaluar el desempeño a diferentes niveles del rasgo latente** como otra de las principales limitaciones del test. Esto, a su vez, implicaría que únicamente un rango limitado de la población pueda ser testada y que los resultados no se puedan generalizar a la población (Thomson y Oppenheimer, 2016; Toplak et al., 2014). A este respecto, Primi et al. (2015) llevaron a cabo un análisis de los 3 ítems del TRC, de acuerdo con la Teoría de Respuesta al Ítem, y encontraron que estos ítems no son capaces de diferenciar entre un nivel de rasgo latente bajo y medio bajo, así como entre un nivel de rasgo latente medio-alto y alto. Los datos sugieren que el TRC original puede ser un instrumento válido para discriminar entre sujetos con un nivel de RC alto o bajo, pero deficiente para discriminar entre niveles de rasgo más próximos o similares.

Del mismo modo, la distribución no normal de las puntuaciones podría resultar en una **desventaja cuando se evalúa el efecto lineal de las puntuaciones** del TRC sobre otras medidas. Salgado et al. (2019) indican que la distribución sesgada de las puntuaciones del TRC puede ser

⁴ Para conocer más detalles sobre el procedimiento y cálculo de la integración véase el Apéndice A.

una cuestión problemática en el uso de clásicos análisis estadísticos bivariados y multivariados, así como en la estimación de algunas propiedades psicométricas como, por ejemplo, el coeficiente de fiabilidad por consistencia interna, donde se debe asumir la distribución normal de las puntuaciones. Weller et al. (2013) indican que esta característica ha forzado a los investigadores a usar diseños de grupos divididos en función de la puntuación alcanzada en el TRC (altos vs. bajos en RC) (véase, por ejemplo, Altmann y Falk, 2009; Feltz y Cokely, 2008; Frederick, 2005; Howarth, Handley y Walsh, 2016; Kuhn, Kuhn y Villeval, 2017; Lohse, 2016; Noori, 2016; Reuben, Sapienza y Zingales, 2015).

Asimismo, y a pesar del generalizado uso del TRC, sus propiedades psicométricas no han sido muy investigadas. La fiabilidad del test ha sido otra de las propiedades a las que no se le ha dado la importancia que merece, a pesar de que diversos artículos han estimado y reportado la magnitud de la fiabilidad (Aktaş et al., 2017; Drummond y Fischhoff, 2017; Juanchich et al., 2016; Reyna y Wilhelms, 2017; Skagerlund et al., 2018). El Apéndice A recoge la distribución de los valores de fiabilidad por consistencia interna de diversos estudios primarios. Como se puede observar, los valores de fiabilidad oscilan entre .39 (Teovanović, Knežević, Stankov, 2015) y .83 (Spadaccini y Esteves, 2014). Existe variabilidad entre estos valores ($SD_{r_{xx}} = .088$) y en su gran mayoría los **coeficientes de fiabilidad no** llegan a ser **aceptables**. Por ejemplo, Teovanović et al. (2015) estimaron en .39 la fiabilidad del TRC para una muestra de 243 sujetos. Esto indica que solo el 39% de la varianza observada está determinada por la varianza verdadera, el resto de varianza (61%) es espuria, es decir, está determinada por el error de la medida.

Existe la creencia de que la validez predictiva de una medida es directamente proporcional a su fiabilidad, hasta cierto punto y bajo ciertas condiciones (Guilford y Fruchter, 1984). De modo que cuanto más fiable sea la prueba más válida será esta. Por lo tanto, necesitamos que un test sea fiable para que también pueda ser válido. Las evidencias muestran la falta de fiabilidad del TRC original, lo que necesariamente implica que nos cuestionemos la validez de la medida. Con el objetivo de conocer la fiabilidad promedio del TRC original hemos integrado los 41 coeficientes de fiabilidad por consistencia interna reportados en los estudios empíricos revisados. La Tabla A2 del Apéndice A recoge el valor estimado. Como se puede observar, el coeficiente de fiabilidad promedio estimado fue de .66, por lo que se puede indicar que la fiabilidad del test es otra importante limitación del TRC. Convencionalmente, se ha requerido que la magnitud de los coeficientes de fiabilidad sea la más alta posible. Aunque esto depende de las consecuencias derivadas del uso de las puntuaciones. Así, cuando el instrumento se emplee con la finalidad de tomar decisiones que afecten a las personas, la magnitud del coeficiente de fiabilidad que se exija debería ser alto (.90, por ejemplo). Sin embargo, cuando no se deriven consecuencias importantes para las personas por el uso del test, coeficientes de fiabilidad en torno a .70 pueden ser aceptados (Prieto y Delgado, 2010).

Finalmente, un significativo inconveniente del test ha surgido a lo largo de los años con el **popular uso de esta medida**. Los tres ítems originales han sido objeto de estudio en diversos seminarios y sesiones docentes, se han tratado en cursos y conferencias y forma parte del contenido del libro de Kahneman “Pensar despacio, pensar rápido” (2011), lo que ha provocado un creciente éxito de los problemas, especialmente del ítem del bate y la pelota (Toplak et al., 2014). Consecuentemente, el TRC se ha convertido en un test popular y ampliamente conocido por la gente, con la consiguiente posible **disminución de su eficacia predictiva** (Baron et al., 2015; Brañas-Garza et al., 2015).

En este sentido, en el Apéndice A se puede observar cómo, en tendencia general, las medias de las puntuaciones en el TRC son más altas en los estudios más recientes (últimas filas de la tabla) y más bajas en los primeros estudios llevados a cabo (primeras filas de la tabla). Varios estudios han examinado esta cuestión. Los resultados del meta-análisis de Brañas-Garza et al. (2015) mostraron que, aunque el número de años expuestos al TRC tuvo un impacto sobre el número de ítems correctamente respondidos incrementando el desempeño en el test, estos resultados no fueron significativos. En un estudio primario posterior, Alós-Ferrer et al. (2016) hallaron que los participantes con más conocimiento previo sobre los ítems del TRC significativamente obtuvieron mejores resultados y disminuyeron el tiempo invertido en responder a los ítems. Thomson y Oppenheimer (2016) testaron esta misma cuestión con una muestra de sujetos reclutados por la web *Mechanical Turk* de Amazon (esta plataforma de *crowdsourcing* ha sido, previamente, señalada como una fuente de participantes altamente contaminada por el grado de exposición previa a los ítems del TRC). Los resultados mostraron que los sujetos que indicaron haber estado expuestos a los ítems respondieron significativamente más problemas correctamente que aquellos que no estuvieron previamente expuestos a ellos. Por lo tanto, tener experiencia previa en el desempeño en el test podría inflar las puntuaciones observadas y debilitar la validez o precisión predictiva de la medida.

En su conjunto, las diversas limitaciones del TRC han provocado el desarrollo de nuevos ítems y test de RC (Borghans y Golsteyn, 2014; Finucane y Gullion, 2010; Morsanyi et al., 2014; Primi et al., 2015; Thomson y Oppenheimer, 2016; Toplak et al., 2014). En el Apéndice B se recoge una relación de los nuevos ítems desarrollados y utilizados por otros autores. Como se puede observar, en la mayoría de los casos, se tratan de nuevas versiones de los ítems originales. El Apéndice C recoge los estudios empíricos llevados a cabo con las nuevas versiones, la distribución de las puntuaciones para cada muestra y los coeficientes de fiabilidad reportados, según cada caso. En tendencia general, se puede observar que las nuevas medidas adolecen, nuevamente, de un número reducido de ítems. El número de ítems que las conforman oscilan entre 4 y 13 y en todos los casos los 3 ítems originales están incluidos. En algunas muestras las distribuciones siguen siendo sesgadas, con medias bajas y desviaciones típicas de la misma magnitud o superior que las medias (Calvillo y Burgeno, 2015; Finucane y Gullion, 2010; Toplak

et al., 2014). Sin embargo, en otros casos la distribución de las puntuaciones se aproxima a una distribución normal, como ocurre en el TRC de 11 ítems de Toplak et al. (2017) o en el TRC de 13 ítems de Salgado (2014a). Además, las nuevas versiones del TRC, frecuentemente, presentan bajos coeficientes de fiabilidad (Calvillo y Burgeno, 2015; Morsanyi, McCormack y O'Mahony, 2017; Thomson y Oppenheimer, 2016; Yilmaz y Saribay, 2017). Aunque, en términos generales, el coeficiente de fiabilidad promedio estimado para las nuevas versiones es aceptable en términos psicométricos ($\bar{r}_{xx} = .74$).

Apenas se han investigado otras propiedades psicométricas de las nuevas versiones del TRC. Sin embargo, existe alguna excepción. Primi et al. (2015) examinaron las propiedades de los ítems de su nueva versión de acuerdo con la Teoría de Respuesta al Ítem y encontraron que los nuevos ítems muestran mayores niveles de discriminación que los ítems originales. Además, proporcionan más información alrededor del nivel medio del rasgo, aunque la información es relativamente consistente a lo largo de un amplio rango del rasgo latente. En general, la nueva versión evalúa mejor las diferencias individuales en el constructo de RC.

Por último, las nuevas versiones y el TRC original presentan correlaciones considerables-altas entre sí, de acuerdo con los criterios de Pearson (1904). Por ejemplo, Chandler, Mueller y Paolacci (2014) encontraron que el TRC original y la nueva versión de Finucane y Gullion (2010; compuesta por 3 ítems originales y 3 ítems nuevos) correlacionaron en .79 ($N = 98$). West et al. (2012) indicaron que la correlación entre el TRC original y 7 nuevos ítems fue de .83. Toplak et al. (2014) encontraron una correlación de .86 entre el TRC original y la nueva versión de 7 ítems (3 originales y 4 nuevos desarrollados por los autores; $N = 160$). Asimismo, encontraron una correlación de .58 entre los ítems originales y los 4 nuevos y una correlación de .92 entre la nueva versión de 7 ítems y los 4 nuevos. Por su parte, Thomson y Oppenheimer (2016) hallaron una relación entre el TRC original y la nueva versión verbal de 4 ítems de $r(136) = .511$ ($p < .001$) y cuando se corrigió esta correlación por error en la medida ascendió a .905. Yilmaz y Saribay (2017) encontraron la misma asociación entre ambos test ($r(426) = .51$; $p < .01$). Por último, Primi, Morsanyi, Donati, Galli y Chiesi (2017) encontraron que la nueva versión de 6 ítems se asoció en .93 con el TRC original ($p < .001$; $N = 560$). Los resultados reflejan que los nuevos test parecen compartir una importante varianza explicada con el TRC original y, por lo tanto, el mismo constructo está siendo evaluado con los diversos instrumentos de RC.

A pesar de las mejoras psicométricas encontradas en las nuevas versiones del TRC, no todos los nuevos test han conseguido superar las debilidades expuestas en este apartado. Por lo tanto, todavía parece necesario el desarrollo de un nuevo test de RC que sí consiga hacerlo o la aplicación de las nuevas versiones que ya lo hagan. Parece conveniente, por tanto, utilizar medidas con un número suficiente de ítems que englobe problemas de diferente dificultad, que permitan la evaluación del nivel del rasgo latente a todo el rango de población y que no produzca “efecto suelo”. Asimismo, sería conveniente que la distribución de las puntuaciones se haga de forma

normal o próxima a ella, donde aproximadamente dos tercios de la población obtengan puntuaciones entorno a una desviación típica a cada lado de la media, mientras que un tercio de ella se distribuya de forma equitativa hacia las dos colas de la distribución. De este modo, las puntuaciones se distribuirían de forma simétrica. Además, el desarrollo de escalas más extensas ayudaría a alcanzar una mayor consistencia interna entre los ítems, permitiendo que el test pueda medir de forma fiable el constructo de RC, propiedad necesaria para una adecuada validez de la medida.



CAPÍTULO 3.

Evidencias de Validez de Constructo y Criterio de la Medida de Reflexibilidad Cognitiva (RC)

1. INTRODUCCIÓN

El principal objetivo de este capítulo ha sido sintetizar las evidencias de validez de constructo y criterio de la medida de Reflexibilidad Cognitiva (RC). Los *Standards for Educational and Psychological Testing* (2014) definen la validez como “el grado en que la evidencia y la teoría apoyan la interpretación de las puntuaciones del test de acuerdo con el propósito para el que fue diseñado” (pág. 9). De modo que, la validez es la consideración más relevante en el desarrollo y evaluación de un test. Existen diferentes fuentes que aportan información y evidencias de validez. Por ejemplo, las evidencias basadas en información sobre el contenido del test (palabras, formato de los ítems, tareas que conforman el test), basadas en su estructura interna, en el proceso de respuesta y en la relación con otras variables y criterios. La validación de un test consiste, por tanto, en acumular evidencias de validez procedentes de todas estas fuentes.

Con el objetivo de presentar las evidencias de validez más relevantes del TRC, este capítulo se estructurará del siguiente modo. En primer lugar, se presenta un apartado dedicado a mostrar las evidencias de validez de constructo. Para ello, se sintetizan las principales hipótesis sobre qué mide realmente el TRC y, seguidamente, se resume la relación existente entre este constructo y otras variables de interés. En segundo lugar, se presenta un apartado dedicado a examinar la validez de criterio del TRC. Con este propósito, se sintetizan las evidencias de validez predictiva para explicar criterios académicos y ocupacionales de interés.

2. EVIDENCIAS DE VALIDEZ DE CONSTRUCTO

En este apartado se pretende comprender qué mide realmente el TRC y para ello se aportan evidencias de validez basadas en su relación convergente y discriminante con otras variables.

La evidencia empírica ha establecido tres posibles hipótesis o conjeturas. La primera sostiene que el TRC es una medida de capacidad cognitiva y más concretamente, una medida de capacidad cognitiva numérica. Los autores que defienden esta hipótesis sostienen que la capacidad numérica conduce la capacidad del TRC para predecir mejores decisiones (Sinayev y Peters, 2015; Weller et al., 2013). A este respecto, varios autores han demostrado la existencia de una relación positiva entre el TRC y diversas medidas de capacidad cognitiva (Frederick, 2005), siendo las medidas de capacidad matemática las que más robustamente se han asociado con el desempeño en el TRC. No obstante, también se ha mostrado que el TRC se asocia con tareas y con comportamientos de toma de decisión que no implican capacidad numérica o matemática y que, en ocasiones, el TRC explica varianza adicional de esas medidas respecto de los test de capacidad cognitiva o matemática (Primi et al., 2015; Shenhav, Rand y Greene, 2017). Estos resultados sugieren que la capacidad cognitiva no es el único determinante del desempeño en el TRC y que este test mide algo más que capacidad numérica (Liberali et al., 2012). Así, dos nuevas hipótesis han sido propuestas.

Algunos autores sostienen que se trata de un test que mide capacidad o disposición hacia un pensamiento racional, es decir, es la capacidad para inhibir el procesamiento avaro o miserable de nuestro sistema 1 y la tendencia a activar el procesamiento deliberado del sistema 2 (Frederick, 2005; Toplak et al., 2011, 2014). Toplak y colaboradores (2011) defienden que los seres humanos somos tacaños cognitivos por defecto, en el sentido de que tendemos a emplear atajos cognitivos cuando pensamos o procesamos información. De este modo, el TRC nos permite medir la capacidad o disposición de los sujetos a inhibir conscientemente este procesamiento avaro a favor de un procesamiento más racional. No obstante, otros autores defienden que la capacidad cognitiva es importante para el correcto desempeño en el TRC pero que, además, este mide la disposición a emplear un procesamiento racional de mente abierta. Es decir, el TRC es, además, una disposición de pensamiento minuciosos, reflexivo y de cognición elaborada. Identifican la RC con la noción de *reflectiveness* o de *actively open-minded thinking* (AOT) de Baron (1988) e, incluso, defienden que el TRC mide, además, la capacidad de las personas para determinar el estilo de pensamiento más apropiado a cada situación (Campitelli y Labollita, 2010; Cokely and Kelley, 2009).

A continuación, se presenta un examen más detallado de las tres hipótesis propuestas, hasta el momento, en la literatura científica.

2.1. ¿Es el TRC una Medida de Capacidad Cognitiva?

En el desarrollo del TRC, Frederick (2005; Kahneman y Frederick, 2002) mostró una asociación entre las puntuaciones del TRC y la capacidad cognitiva de los sujetos. Concretamente,

encontró una relación con la medida *Wonderlic Personnel Test* (WPT) de .43 ($n = 921$). Asimismo, Frederick (2005) examinó la relación entre el TRC y otras medidas de inteligencia como son los exámenes estandarizados de entrada a estudios universitarios, el *Scholastic Aptitude Test* (SAT) y el *American College Testing* (ACT). Ambas pruebas correlacionaron con el test en .44 y .46, respectivamente. De forma desglosada, la correlación entre el TRC y SAT matemático fue de .46 y la relación entre el TRC y SAT verbal de .24.

Desde la publicación de este estudio, sucesivos artículos han reflejado asociaciones similares entre el TRC y múltiples medidas de capacidad cognitiva general (Cokely et al., 2012; Duttie, 2015; Goldwater, Don, Krusche y Livesey, 2018; Hanaki Jacquement, Luchini y Zylbersztejn, 2016; Jackson, Kleitman, Howie y Stankov, 2016; Primi et al., 2015; Ruffle y Tobol, 2017). Así, por ejemplo, Liberali y colaboradores (2012) y Teovanović et al. (2015) demostraron una relación de .31 ($n = 190$) y .25 ($n = 243$) con el test abreviado de *Raven Advanced Progressive Matrices*, respectivamente. Brañas-Garza et al. (2012) examinaron la relación entre ese mismo test con los diversos ítems del TRC por separado. Los resultados mostraron una correlación con el ítem 1 del TRC de .16 ($p = .025$), con el ítem 2 de .08 ($p = .259$) y con el ítem 3 de .32 ($p < .001$). Por su parte, Thompson et al. (2013) encontraron una asociación con el test de IQ *Shipley Institute of Living Scale*, ($r(142) = .43$; $p < .001$) y Toplak et al. (2014) mostraron una correlación positiva y significativa con la medida de *Wechsler Abbreviated Scales of Intelligence* ($r(160) = .48$ y $r(160) = .50$ para el TRC original y el desarrollado por los autores, respectivamente). Los resultados de Fehr y Huck (2015) y de Narayanan y Moritz (2015) reflejaron correlaciones moderadas con el WPT. En ambos casos, los tamaños del efecto fueron similares al encontrado por Frederick (2005). Finalmente, Morsanyi et al. (2017) hallaron una asociación con el test de *ordering abilities* de $r = .45$ ($n = 69$).

Varios estudios se han centrado en examinar la relación entre el TRC y el SAT. Los resultados del experimento de Thompson et al. (2013) reflejaron una correlación con la medida de SAT de .34, ($p < .01$; $n = 150$). Insler, Compton y Schmitt (2015) hallaron una asociación de .39 ($n = 364$) y Thomson y Oppenheimer (2016) encontraron una relación con SAT auto informado de $r = .59$ ($n = 134$). Del mismo modo, algunos estudios han indicado la relación del TRC con las diversas pruebas de SAT. Así, por ejemplo, Obrecht et al. (2009) hallaron que los sujetos con altas puntuaciones en el TRC obtienen simultáneamente puntuaciones más altas en SAT matemático ($r(417) = .45$). Poore, Forlines, Miller, Regan e Irvine (2014) también demostraron una relación con SAT matemático ($r(459) = .478$, $p < .10$) y con SAT verbal ($r(456) = .219$; $p < .10$). Por último, los resultados de Ventis (2015) mostraron que el TRC correlacionó significativamente con SAT matemático ($r(148) = .43$), con SAT de lectura crítica ($r(148) = .28$) y con SAT escrito ($r(148) = .35$).

Todos estos resultados sugieren la existencia de una relación moderada entre el TRC y la capacidad cognitiva general, de acuerdo con los criterios de Pearson (1904), lo que, a su vez, muestra que todas ellas reflejan factores comunes. No obstante, aunque los constructos comparten varianza, el tamaño de las correlaciones sugiere que las variables podrían tener una fuente en común pero no capturar la misma capacidad cognitiva (Hanaki et al., 2016; Liberali et al., 2012). Podrían, por tanto, medir características distintas en función del objetivo para el que fueron diseñadas (Frederick, 2005). Así, los test de inteligencia miden la capacidad cognitiva general de los sujetos, es decir, la habilidad o capacidad para aprender y adquirir nuevos conocimientos y destrezas (Cattell y Cattell, 1986; Hunter, 1986; Salgado y Moscoso, 2008). Mientras que las pruebas de ACT y SAT miden el logro o desempeño académico y el TRC mide la RC, es decir, la capacidad o disposición a resistirse a responder con la respuesta que primero viene a la mente (Frederik, 2005).

En este mismo sentido, Guthrie et al. (2007) indicaron que, aunque el TRC actúa como un test de IQ abreviado que mide algunos componentes de inteligencia, pensar en el TRC como una simple medida de capacidad mental general podría ser un error. Para estos autores, el TRC mide la capacidad o disposición a resolver problemas deliberadamente cuando la intuición podría inducir a error (definición que tradicionalmente los psicólogos incluyeron en el constructo de inteligencia). Lohse (2016) indica que el TRC no es una medida de inteligencia general, sino que representa una subcategoría específica de un amplio conjunto de capacidades cognitivas. Así, Stanovich (2011) diferencia entre los conceptos de inteligencia y racionalidad y afirma que deficiencias cognitivas en la toma de decisión racional no necesariamente implicaría peores puntuaciones en el desempeño de los test de inteligencia. Acuñó el término *dysrationalia* (dislexia) para referirse a la incapacidad de pensar y comportarse racionalmente a pesar de tener una adecuada inteligencia. En relación con esto, Stanovich (2011) sostiene que la tradicional definición de inteligencia engloba características propias de un pensamiento racional. Sin embargo, los test estándar de IQ no miden esas características, sino simplemente un pequeño subconjunto de capacidades de pensamiento. Por ejemplo, Stanovich (2011) defiende que pensar de forma racional implica adoptar objetivos apropiados, tomar cursos de acción congruentes con nuestros objetivos y creencias y defender las creencias que son acordes con las evidencias disponibles. Asimismo, indica que, aunque los test de IQ evalúan la capacidad para enfocarse a un objetivo inmediato sin permitir distracciones, no evalúan si los objetivos adoptados por las personas son racionales en un primer momento. Del mismo modo, los test de inteligencia miden la capacidad de los sujetos para mantener y manipular las creencias en la memoria a corto plazo, pero no evalúan si las personas muestran una tendencia a formar creencias racionales cuando se presentan evidencias. Por último, los test de IQ evalúan si las personas son eficientes en el

procesamiento de la información que le es facilitada, pero no examinan si las personas son evaluadores críticos de la información ofrecida por el entorno.

Corgnet, Espín y Hernán-González (2015) defienden que el TRC se diferencia de inteligencia por no ser una medida estándar de IQ. Los test de IQ miden la capacidad para resolver problemas, pero fallan a la hora de medir la capacidad para activar un procesamiento reflexivo. De modo que los autores sostienen que el TRC no solo mide IQ sino también RC.

Sin embargo, diversos autores defienden que las puntuaciones del TRC podrían depender de la capacidad de controlar y revisar el contenido de la memoria operativa y, por ende, las diferencias individuales en la toma de decisiones frecuentemente son debidas a las limitaciones de procesamiento de la memoria operativa (Graffeo et al., 2015; Stanovich y West, 2000). A este respecto, Welsh et al. (2013) indican que la inmensa mayoría de artículos que han examinado la relación entre el TRC e inteligencia han utilizado medidas de capacidad mental general (CMG) o de capacidad numérica obviando los diversos tipos de capacidades cognitivas que el modelo de Cattell-Horn-Carroll propuso. Sostienen que es posible que se haya subestimado la relevancia de específicas capacidades cognitivas en la predicción de un pensamiento racional. Actualmente, existen evidencias que muestran que las puntuaciones del TRC se relacionan con diversas medidas de memoria operativa como, por ejemplo, *Working Memory Capacity Span* (Cokely y Kelley, 2009; Stuppel Gale y Richmond, 2013), *Backward Digit Span* (Finucane y Gullion, 2010), *Paired Associates* (Finucane y Gullion, 2010), *Semantic Anaphora Working Memory Test* (Gómez-Chacón et al., 2014) y *O-span* (Goldwater et al., 2018).

Por otro lado, las evidencias empíricas también han reflejado una tendencia a que las puntuaciones del TRC se relacionan con la capacidad cognitiva verbal (Jackson et al., 2016; Shenhav et al., 2011; Pennycook et al., 2012; Toplak et al., 2014; Trippas, Pennycook, Verde y Handley, 2015).

Del mismo modo, diversos estudios han evidenciado la asociación entre el TRC y otras medidas cognitivas. Por ejemplo, Finucane y Gullion (2010) hallaron una correlación entre el TRC y medidas de velocidad perceptual (*Letter Comparisson Test*: $r(219) = .23$ y *Digit Symbol Substitution Test*: $r(159) = .24$). Poore et al. (2014) encontraron una relación con los resultados en problemas de analogías ($r(709) = -.105$) y en problemas de lógica ($r(695) = .348$). Teovanović et al. (2015) hallaron una asociación con medidas de capacidad mecánico-espacial ($r(243) = .35$). Koscielniak et al. (2016) hallaron una asociación con velocidad mental medido con *Digit Symbol Substitution Test* ($r(158) = .210$) y, por último, Lindeman y Svedholm-Häkkinen (2016) mostraron una relación con rotación mental ($r(258) = .19$).

El TRC también se asoció con medidas de capacidad numérica o matemática. En este sentido, se ha encontrado una correlación con fluidez matemática ($r(59) = .28$ y $r(59) = .42$ para

el TRC original y el TRC desarrollado, respectivamente, Primi et al., 2015 y $r(69) = .40$, Morsanyi et al., 2017), con resolución de problemas matemáticos ($r(431) = .498$; Poore et al., 2014), con la realización de sumas ($r(264) = .26$, Corgnet, Hernán-González y Mateo, 2015) y con Numerical Ability Test ($r(102) = .44$, Welsh et al., 2013).

Sin embargo, la medida de capacidad matemática más ampliamente examinada en la literatura del TRC ha sido las Numeracy Scales. Graffeo et al. (2015) definen *numeracy* como “la capacidad de comprender, usar y adjuntar el significado de los números” (pág. 2). Incluye la capacidad de componer magnitudes y comprender conceptos en ratios (fracciones, proporciones, porcentajes y probabilidades). Ghazal, Cokely y Garcia-Retamero (2014) afirman que existe un amplio consenso en que el constructo teórico de *numeracy* no es sinónimo de una habilidad matemática pura, pero sí se refiere a la literatura matemática o cuantitativa. Concretamente, el constructo *numeracy* se relaciona con las competencias matemáticas que son necesarias en nuestra vida cotidiana, que están conectadas con el contexto, el propósito o el uso. A nivel más básico, *numeracy* envuelve números reales, tiempo, medidas y estimación; mientras que, a altos niveles, *numeracy* se enfoca en entender conceptos que implican ratios, fracciones, proporciones, porcentajes y probabilidades (Ghazal et al., 2014). Así, se ha encontrado que las puntuaciones en el TRC se relacionan con la escala de 3 ítems de *numeracy* de Schwartz, Woloshin, Black y Welch (1997; $r(608) = .53$, Finucane y Gullion, 2010), con la escala de *numeracy* de 10 ítems desarrollada por Lipkus, Samsa y Rimer (2001; $r(69) = .31$, Cokely y Kelley, 2009), con la escala de 11 ítems de *numeracy* de Peters y Levin (2008; $r(259) = .51$ y $r(190) = .39$, para el experimento 1 y 2, respectivamente, Liberali et al., 2012), con *Berlin Numeracy Test* de Cokely et al. (2012; $r(300) = .51$, $r(300) = .56$ y $r(300) = .41$, para el formato adaptativo de ordenador, papel y lápiz y formato abreviado de un único ítem, respectivamente), con *Rasch-based Numeracy Scale* de Weller et al. (2013; $r(1409) = .57$, Sinayev y Peters, 2015) y con la escala de 7 ítems de *numeracy* ($r(2063) = .61$; Skagerlund et al., 2018).

Diversos estudios también han examinado la relación entre el TRC y otras medidas de competencia matemática o numérica que están más íntimamente ligadas al contexto de nuestra vida cotidiana. Así, se ha encontrado que las puntuaciones del TRC correlacionan con el razonamiento probabilístico ($r(560) = .30$, Primi et al., 2017), con tareas de probabilidad condicional ($r(226) = .22$, Primi et al., 2015), con razonamiento transitivo ($r(69) = .26$, Morsanyi et al., 2017), con razonamiento deductivo condicional ($r(69) = .38$, Morsanyi et al., 2017), con el compuesto de las medidas de razonamiento probabilístico y la escala de Berlin Numeracy Test ($r(181) = .59$; Primi et al., 2018), con el razonamiento científico ($r(395) = .38$, $p < .001$, en el estudio 1c y $r(270) = .45$, $p < .001$ en el estudio 2; Drummond y Fischhoff, 2017) y con el razonamiento bayesiano ($p = .254$, $p < .001$, Lesage et al., 2013). Los datos sugieren que los

sujetos con altas puntuaciones en el TRC están más expuestos a datos estadísticos en general (Obrecht et al., 2009).

Asimismo, se ha demostrado una relación entre el TRC y algunas medidas de capacidad matemática subjetiva. Por ejemplo, se ha mostrado una correlación moderada entre el TRC y la escala subjetiva de *numeracy* (Gómez-Chacón et al., 2014; Liberali et al., 2012; Poore et al., 2014; Primi et al., 2015, 2017). Las medidas subjetivas de *numeracy* fueron diseñadas con el objetivo de evaluar lo seguro y cómodo que la gente se siente con su capacidad para entender y aplicar los números sin estar, en ese momento, realizando o desempeñando tareas numéricas. Los estudios reflejaron tamaños del efecto en torno a .19 y .48. Por su parte, Gómez-Chacón et al. (2014) mostraron que el TRC correlacionó significativamente con el sistema de creencias sobre las matemáticas ($r(56) = .341$).

En su conjunto, todos estos resultados sugieren que las puntuaciones del TRC correlacionan tanto con inteligencia cristalizada como con inteligencia fluida (Corgnet, Espín y Hernán-González, 2015; Finucane y Gullion, 2010) y con una amplia variedad de medidas de capacidad cognitiva. No obstante, se puede observar que el TRC presenta tamaños del efecto más robustos cuando se asocia con capacidad numérica que cuando lo hace con otras medidas de capacidad cognitiva. Estos datos apoyan la hipótesis de que el TRC es una medida más de capacidad numérica. En la literatura científica existen, incluso, especulaciones sobre la existencia de una subyacente dimensión numérica en el TRC (Weller et al., 2013).

A este respecto, Finucane y Gullion (2010) llevaron a cabo un análisis de regresión para determinar las puntuaciones en el TRC y encontraron que las capacidades cognitivas, en su conjunto, explicaron el 13.2% de la varianza de la RC, mientras que, la medida de *numeracy*, individualmente, explicó el 10.4%. Después de controlar por las restantes variables de capacidad cognitiva, *numeracy* explicó un 4.5% adicional de la varianza del TRC. Por lo que, esta medida explicó un 44.7% de la varianza de la RC relacionada con las capacidades cognitivas.

Weller et al. (2013) desarrollaron una nueva escala de *numeracy* denominada *Rasch-based Numeracy Scale*. Para ello, llevaron a cabo dos análisis factoriales con las escalas de *numeracy* de 3 ítems de Schwartz et al. (1997), la escala de 8 ítems de Lipkus et al. (2001) y el desarrollo de esta última con la incorporación de 4 ítems diseñados por Peters, Hibbard, Slovic y Dieckmann (2007). Asimismo, incluyeron el TRC para explorar su ajuste como posible medida del constructo de *numeracy*. Los resultados de ambos análisis factoriales confirmatorios mostraron un ajuste estadístico idéntico para el modelo de 1 factor (*numeracy*) o de dos factores (*numeracy* y RC). El modelo de dos factores no ofreció un mejor ajuste que el modelo de 1 factor. Además, la correlación entre ambos factores fue alta, lo que sugirió la existencia de un único

factor. El segundo estudio de Weller et al. (2013) replicó los resultados obtenidos en el primero. Los datos mostraron la misma estructura factorial.

Liberali et al. (2012) llevaron a cabo un análisis factorial confirmatorio con el objetivo de examinar si diversas medidas de *numeracy* y el TRC miden lo mismo. Los resultados mostraron que las medidas cargaron hacia 6 factores diferentes. Dos de ellos, se correspondieron con las dos dimensiones en que se compone la medida subjetiva de *numeracy*, otro recogió los ítems del TRC y los tres ítems de la escala general de *numeracy* (Schwartz et al., 1997) y los tres factores restantes hicieron referencia a diferentes dimensiones de la escala de *numeracy* de 8 ítems de Lipkus et al. (2001; multiplicaciones, proporciones y significado de las magnitudes). De modo que, los resultados sugirieron, una vez más, que el TRC es un constructo homólogo de las medidas de *numeracy*. Thompson, Pennycook, Trippas y Evans (2018) encontraron resultados en esta misma línea. El test de IQ general, el TRC, la escala de *numeracy* y los problemas de tasa base cargaron hacia el mismo factor. Sin embargo, el análisis de regresión del experimento 2 de Liberali et al. (2012) reflejó que el TRC y la medida general de *numeracy* cargaron hacia factores separados, sugiriendo que, aunque puedan estar relacionadas entre sí, pueden tratarse de medidas independientes. Los resultados de este último experimento apoyan la hipótesis de que el TRC no es una medida más de *numeracy* o de conocimientos básicos de aritmética.

Por lo tanto, aunque análisis factoriales confirmatorios previos han sugerido que los constructos de RC y capacidad numérica pueden ser considerados un único factor, Ghazal et al. (2014) afirmaron que hay razones para ser precavidos con esta interpretación por dos motivos: (1) en ciertos casos, el TRC predice los resultados de la toma de decisión más allá de *numeracy* (Cokely et al., 2012; Moritz et al., 2014; Toplak et al., 2011) y (2) estudios previos también han demostrado que ambos tipos de ítems pueden cargar hacia factores diferentes (Liberali et al., 2012). Juanchich et al. (2016) añadieron tres motivos más: (1) *numeracy* no media el efecto del TRC en la capacidad de razonamiento (Sirota y Juanchich, 2011); (2) el TRC predice el desempeño de tareas que no requieren cálculo matemático (Campitelli y Labollita, 2010; Königsheim et al., 2018) y (3) el desempeño en el TRC no puede ser completamente explicado por la capacidad numérica (Campitelli y Gerrans, 2014). En relación con este último punto, Moritz et al. (2014) encontraron que los individuos con bajas puntuaciones en capacidad cognitiva también muestran bajas puntuaciones en el TRC, sin embargo, los sujetos con altas puntuaciones en IQ muestran bajas o altas puntuaciones en el TRC al menos con la misma frecuencia. Este dato apoya la afirmación de que la capacidad cognitiva no puede explicar la totalidad de la RC.

Resumiendo, los estudios empíricos han reflejado una asociación moderada entre el TRC y diversas medidas de capacidad cognitiva, no solo de capacidad mental general, sino también de capacidades cognitivas específicas. Por lo que, el TRC mantiene una relación no solo con

inteligencia fluida sino también con inteligencia cristalizada. No obstante, se ha mostrado que los tamaños del efecto son más robustos entre el TRC y capacidad numérica que entre el TRC y el resto de las medidas cognitivas, lo que apoya la hipótesis de que ambos constructos se refieran en realidad al mismo concepto teórico. No obstante, estudios recientes han encontrado evidencias de que el TRC y las medidas de capacidad cognitiva podrían no referirse exactamente al mismo constructo, ya que predicen de forma muy variada otras medidas de razonamiento, a pesar de su varianza compartida (Liberali et al., 2012; Moritz et al., 2014; Sirota y Juanchich, 2011; Toplak et al., 2011). A este respecto, diversos autores han propuesto la hipótesis de la RC que defiende la idea de que el TRC es una medida de capacidad, entendida como diferentes formas de capacidad cognitiva. Es decir, definen el TRC como la tendencia a comprobar y detectar errores intuitivos (Sinayev y Peters, 2015). Toplak et al. (2011) denominaron a esta capacidad pensamiento racional y proponen que el TRC directamente mide la capacidad de pensamiento racional o la disposición a no procesar por defecto de forma avara o miserable.

2.2. ¿Es el TRC una Medida de Capacidad Matemática y Pensamiento Racional?

Frederick (2005) definió la RC como una *capacidad* o *disposición* para inhibir la respuesta inmediata de nuestra mente y activar los mecanismos de deliberación. La evidencia empírica ha sugerido que los sujetos que obtienen altas puntuaciones en el TRC también suelen alcanzar buenos resultados en los test de capacidad cognitiva, lo que apoya la identificación de la RC como una capacidad. No obstante, la definición del constructo engloba la posibilidad de que la RC sea una disposición. En este sentido, Toplak et al. (2011) definen la RC como la disposición hacia un pensamiento racional, entendido como la tendencia a no procesar de forma miserable para reducir la carga cognitiva requerida en el procesamiento.

La noción de miseria cognitiva fue inicialmente propuesta por Fiske y Taylor (1984) para ilustrar el hecho de que los individuos, comúnmente, evalúan la información y toman decisiones usando atajos cognitivos. Los atajos o sesgos cognitivos se refieren a los errores sistemáticos que la gente comete en acciones de elección, en los juicios y en la estimación de las probabilidades (Toplak et al., 2017). Toplak y colaboradores (2014) identifican el concepto de miseria cognitiva con la tendencia básica de los humanos a emplear un procesamiento de tipo 1. Estudios previos sugieren que esta tendencia podría ser una cuestión biológica de evolución adaptativa, lo que indica que los humanos somos frecuentemente no racionales. El tipo 1 de procesamiento nos ofrece una solución rápida que es una primera aproximación a la respuesta óptima, pero llevar a cabo juicios aproximados podría no ser suficiente de acuerdo con las exigencias del entorno. La vida tecnológica requiere de mayor precisión en nuestros pensamientos, de modo que, procesar de forma tacaña a menudo supone un serio problema para alcanzar nuestros objetivos. De este

modo, Toplak et al. (2011) identifican al TRC como un test diseñado para medir la tendencia a anular la respuesta alternativa predominante que es incorrecta, y activar el procesamiento reflexivo que lidera la respuesta correcta.

En relación con esto, Lesage et al. (2013) definen el TRC como un *proxy* de las funciones metacognitivas de inhibición y monitorización que son directa e indirectamente evaluadas en medidas de funcionamiento ejecutivo (Graffeo et al., 2015; Stanovich, 2011). Los autores afirman que alcanzar la solución correcta requiere de la inhibición de la respuesta impulsiva y que, para ello, se precisa de recursos ejecutivos que permitan activar el procesamiento de tipo 2. De forma que el TRC se asocia con la regulación cognitiva y se encarga de controlar y ajustar el proceso de razonamiento (Gómez-Chacón et al., 2014; Thompson, 2009, 2010; Thompson et al., 2013).

De acuerdo con Johnson et al. (2014), una de las mayores manifestaciones de la necesaria capacidad de monitorización e inhibición para el correcto desempeño de los ítems en el TRC, lo demuestran los estudios que examinan el desempeño del test en la versión conflicto y en la versión de no conflicto. Estos estudios han reflejado que, a pesar de que ambas versiones de los ítems implican los mismos cálculos matemáticos, un mayor porcentaje de sujetos fallan al responder los ítems en la versión original o de conflicto que en la versión modificada o de no conflicto (De Neys et al., 2013; Johnson et al., 2014; Szollosi et al., 2017). Los autores atribuyen las diferencias en el desempeño a la capacidad ejecutiva que los ítems demandan en la versión conflicto. Por lo tanto, sugieren que el TRC no mide únicamente capacidad matemática, sino que, para resolver correctamente el problema, se necesita activar el sistema 2 que requiere, en mayor medida, de los recursos de la capacidad controlada de ejecución. En este sentido, el TRC es un test que captura capacidad cognitiva y disposición hacia un pensamiento analítico.

Del Missier et al. (2011) mostraron la existencia de una relación entre el funcionamiento ejecutivo y la RC. Los autores examinaron la asociación entre las puntuaciones en el TRC y tres roles de funcionamiento ejecutivo: (1) cambio (*shifting*) que hace referencia a la flexibilidad mental; (2) actualización (*updating*) que hace referencia a la capacidad para actualizar las representaciones de la memoria operativa, se refiere, por tanto, a la activa monitorización y revisión de las representaciones de la memoria operativa; y (3) inhibición (*inhibition*) que hace referencia a la capacidad para controlar las interferencias de otros estímulos y mantener la atención enfocada en los objetivos relevantes. El modelo de ecuaciones estructurales reflejó que las funciones de inhibición y actualización se unificaron en un factor único y que se asociaron con el TRC, mientras que el rol de cambio se mantuvo como factor independiente. Los resultados sugieren que el desempeño en el TRC está relacionado con la función de monitorización e inhibición necesaria para detener la respuesta intuitiva y errónea que inicialmente provocan los ítems del TRC.

Asimismo, Del Missier et al. (2011) también llevaron a cabo un análisis de mediación para comprobar si la relación entre la función de monitorización/inhibición y el desempeño en el TRC está mediado por la CMG o capacidad numérica. Los resultados mostraron que el tamaño del efecto directo de la función de monitorización/inhibición sobre el TRC fue de .31 mientras que, el efecto indirecto a través de inteligencia fluida fue de .007 y a través de *numeracy* de .118. Por lo tanto, el efecto indirecto total fue de .13. Los resultados reflejaron que, únicamente, *numeracy* parcialmente medió la relación entre la función de monitorización/inhibición y el desempeño en el test y que, por tanto, la capacidad matemática es un recurso necesario para el correcto desempeño en el test.

Sin embargo, estudios posteriores no hallaron una asociación entre el test y las medidas de inhibición. En el análisis factorial llevado a cabo por Liberali et al. (2012) se encontró que el TRC y las medidas de inhibición cargaron hacia factores diferentes, sugiriendo que, ambas medidas son independientes, aunque puedan estar relacionadas entre sí. Por su parte, Welsh et al. (2013) encontraron que el TRC cargó hacia el mismo factor que las medidas de capacidad cognitiva y *numeracy*, siendo su carga factorial más próxima a la carga factorial de la medida de *numeracy*; mientras que, las medidas de funcionamiento ejecutivo no cargaron hacia ningún factor.

En conjunto, los datos anteriores sugieren que poseer altas capacidades cognitivas no es garantía suficiente de un buen desempeño en el TRC, sino que se requiere de la capacidad o disposición para pensar analíticamente y activar el procesamiento reflexivo que suprima la respuesta intuitiva (Pennycook, Cheyne, Koehler et al., 2015). Por tanto, se ha identificado al TRC como una herramienta construida idealmente para predecir el desempeño en tareas de heurísticos y sesgos cognitivos, es decir, para testar la tendencia de las personas a caer en errores cognitivos (Albaity, Rahman y Shahidul, 2014; Toplak et al., 2011).

En este sentido, diversos estudios han explorado las propiedades predictivas del TRC en un amplio conjunto de heurísticos, indicando que el TRC se relaciona con la comisión de sesgos cognitivos. Así, se ha mostrado que los sujetos con bajas puntuaciones en el TRC son generalmente menos pacientes en el cobro de una hipotética recompensa monetaria que los sujetos que responden correctamente a todos los ítems en el TRC. Es decir, prefieren una gratificación menor e inmediata frente a una gratificación mayor pospuesta en el tiempo (Albaity et al., 2014; Bosch-Domènech, Brañas-Garza y Espín, 2014; Frederick, 2005; Insler et al., 2013, 2015; Kuhn et al., 2017; Primi et al., 2015; Reuben et al., 2015; Reyna y Wilhelms, 2017; Shenhav et al., 2017; van der Heijden, Klein, Müller y Potters, 2012). No obstante, se ha encontrado que el TRC no predice las preferencias temporales para eventos que impliquen salud, presumiblemente porque estas preferencias son relativamente más afectivas (Frederick, 2005; Hardisty y Weber,

2009). Asimismo, los sujetos con altas puntuaciones en el TRC presentan una menor aversión al riesgo en el dominio de las ganancias que los sujetos con puntuaciones bajas en el TRC. De modo que, escogen opciones más arriesgadas en escenarios hipotéticos, incluso cuando el valor esperado de la opción que implica riesgo es más bajo que el valor esperado de la opción que implica certeza (Campitelli y Labollita, 2010; Cueva et al., 2015; Ferrara, Bottasso, Tempesta, Carrieri, De Gennaro y Ponti, 2015; Frederick, 2005; Insler et al., 2013, 2015; Noori, 2016; Simonovic et al., 2017; Sinayev y Peters, 2015; Taylor, 2016). Sin embargo, en el dominio de las pérdidas no se han encontrado resultados estadísticamente significativos (Campitelli y Labollita, 2010; Thomson y Oppenheimer, 2016). Paradójicamente, se ha hallado que los sujetos con puntuaciones altas en el TRC tienden a mostrar una menor preferencia por la ambigüedad de riesgo (Cesarini, Johannesson, Magnusson y Wallace, 2012; Fairley y Weitzel, 2017), aunque son más tolerantes en sí a la ambigüedad de la información (Beatty y Thompson, 2012; Koscielniak et al., 2016; Poore et al., 2014; Ventis, 2015). Además, algunos estudios han sugerido que la RC podría no asociarse con la preferencia al riesgo en escenarios reales de juego o apuestas (Taylor, 2016), con comportamientos de riesgo en el consumo de alcohol en adolescentes (Reyna y Wilhelms, 2017) ni con acciones criminales de la vida real (Reyna et al., 2018).

Por otro lado, se ha confirmado que los sujetos con baja RC cometen en mayor proporción la falacia de conjunción (Aczel, Bago, Szollosi, Foldes y Lukacs, 2015; Alós-Ferrer et al., 2016; Alós-Ferrer y Hügelschäfer, 2016; Noori, 2016; Oechssler et al., 2009), están más expuestos al sesgo de dar más peso a la información sobre vivencias individuales que a la información probabilística (sesgo de tasa base; Aczel et al., 2015; Hoppe y Kusterer, 2011; Noori, 2016; Pennycook et al., 2014a; Teovanović et al., 2015; Thompson et al., 2018; Trippas et al., 2015; Yilmaz y Saribay, 2017) y al sesgo de conservadurismo (Albaity et al., 2014; Alós-Ferrer y Hügelschäfer, 2016; Hoppe y Kusterer, 2011; Noori, 2016; Obrecht et al., 2009). Ignoran más fácilmente la información del tamaño de la muestra cuando emiten juicios (Aczel et al., 2015; Cesarini et al., 2012; Obrecht et al., 2009), están más contaminados por el estatus social o el estatus referente de las personas para tomar decisiones (Baldi et al., 2013), presentan una mayor tendencia a la comisión del sesgo de creencia en el razonamiento silogístico (Eriksson y Jansson, 2016; Hopkins, Weisberg y Taylor, 2016; Howarth et al., 2016; Szasz et al., 2017; Thomson y Oppenheimer, 2016; Trémoière, De Neys y Bonnefon, 2014; Viator et al., 2014) y son más susceptibles a los juicios sesgados por estereotipos de sexo (Lubian y Untertriffler, 2013).

Sin embargo, las evidencias científicas apuntan a que el TRC no predice el efecto anclaje (Bergman et al., 2010; Oechssler et al., 2009; Teovanović et al., 2015) ni la sobre confianza de los sujetos en su desempeño (*overconfidence*; Hoppe y Kusterer, 2011; Langnickel y Zeisberger, 2016; Noori, 2016; Sinayev y Peters, 2015; Teovanović et al., 2015). Igualmente, parece no existir

un consenso entre si los sujetos más intuitivos están más contaminados por el efecto enmarque que los sujetos más reflexivos (Aczel et al., 2015; Primi et al., 2015; Sinayev y Peters, 2015; Weller et al., 2013) o, por el contrario, la relación entre la RC y la comisión del sesgo es nula (Aczel et al., 2015; Jackson et al., 2016; Toplak et al., 2014). Tampoco se ha podido demostrar una mayor asociación entre los sujetos con puntuaciones bajas en RC y la exposición al efecto del formato probabilístico “1-de-X” (Sirota, Juanchich, Kostopoulou y Hanak, 2014). Para una revisión más detallada de la relación entre la RC y los sesgos cognitivos véase la Tabla 5. Esta tabla recoge una relación de los principales heurísticos examinados, su definición, la relación con el TRC y ejemplos de tamaños del efecto encontrados.

Los resultados evidencian que el TRC es un buen predictor de múltiples sesgos cognitivos, siendo los sujetos con mejores resultados en el TRC los que tienden a ser menos susceptibles a los errores cognitivos. Consecuentemente, la RC es un buen predictor de modelos normativos de juicio y toma de decisión (Sirota y Juanchich, 2011). Asimismo, los resultados expuestos muestran que el TRC correlaciona de forma muy variada con los diversos heurísticos. Esta variabilidad podría ser explicada, en parte, por el componente genético compartido entre ambos constructos. En este sentido, estudios previos han sugerido la existencia de un importante componente hereditario que determina el desempeño en el TRC y la exposición a los atajos cognitivos, de modo que, las correlaciones de las puntuaciones alcanzadas en el TRC y en tareas de heurísticos son más robustas para los gemelos monocigóticos que para los gemelos dicigóticos (Cesarini et al., 2012; Simonson y Sela, 2011). Asimismo, Cesarini et al. (2012) encontraron que la variación heredada en la susceptibilidad a 3 de los 8 sesgos cognitivos examinados correlacionó positivamente con la variación hereditaria en las puntuaciones del TRC.

Del mismo modo, Toplak et al. (2014) afirman que clasificar el TRC como un indicador de pensamiento racional no significa que correlacionará de la misma forma con cada tarea de pensamiento racional, ya que el TRC refleja varianza debido a la capacidad cognitiva algorítmica y, por lo tanto, mostrará variación en la correlación con otras tareas de pensamiento racional en función de la variabilidad de la correlación de estas últimas con las medidas de capacidad cognitiva. En este mismo sentido, Welsh et al. (2013; Campitelli y Labollita, 2010) sugieren que el TRC y varios problemas de toma de decisión suelen sustentarse sobre el cálculo matemático para determinar la respuesta normativa. Esta característica también podría determinar la variabilidad de la relación de la RC con los diversos heurísticos.

Dado el sesgo numérico en la respuesta a los ítems del TRC, los autores defienden que el test podría únicamente explicar la comisión de sesgos cognitivos que envuelven cálculos numéricos o explicar más varianza de estos sesgos en comparación con aquellos que no comprenden cálculo numérico. Los resultados de estudios previos han expuesto evidencias de

casos en los que el TRC predice la exposición a heurísticos que no implican capacidad numérica, como es el caso del sesgo de creencia, la falacia de conjunción o el sesgo de *status quo*, y casos en los que el TRC no predice la comisión de sesgos que sí exigen capacidad matemática, como por ejemplo, el efecto “1-de-X”, la regresión a la media y el sesgo exponencial creciente (para este último caso véase Königsheim et al., 2018). Asimismo, Thomson y Oppenheimer (2016) desarrollaron una nueva escala de RC formada por 4 ítems que no implican cálculo matemático y examinaron su relación con diversas tareas de heurísticos. El análisis de correlaciones reflejó que la relación entre el TRC verbal y los sesgos cognitivos examinados fueron más débiles que las correlaciones entre el TRC original y esos mismos sesgos cognitivos, pero las diferencias entre las correlaciones no fueron estadísticamente significativas. Esto sugiere que la validez del TRC para predecir la comisión de sesgos puede no deberse en sí mismo a la capacidad matemática que envuelve el TRC. De este modo, los resultados reflejan que el TRC no es una medida más de capacidad matemática, sino que se trata de una medida de disposición hacia un pensamiento analítico donde, en efecto, la capacidad matemática es necesaria para hallar la respuesta correcta pero no es el único determinante.

Toplak et al. (2014) sugieren que el poder predictivo del TRC podría ser, en parte, separable de la capacidad cognitiva. El hecho de que el TRC sea una disposición de pensamiento y no tanto una capacidad cognitiva no implica que esta última no sea relevante para el desempeño en el TRC. Por lo que, el TRC incluiría un cierto grado de disposición o propensión a pensar analíticamente además de capacidad cognitiva. Así, Pennycook y Ross (2016) sostienen que, si alguien no cuenta con la disposición a pensar analíticamente, no podrá ejercitar completamente su capacidad cognitiva y no podrá resolver tan bien los problemas del test. Naturalmente, lo contrario también es cierto, si alguien no dispone de capacidad cognitiva suficiente, no importará cuánto tiempo y esfuerzo estén dispuestos a gastar pensando en el problema.

Con el objetivo de comprobar el rol de la RC y otras medidas de capacidad cognitiva en la predicción de los atajos cognitivos se han llevado a cabo diversos estudios. Los resultados han mostrado que el TRC añade validez sobre otras medidas de capacidad cognitiva para predecir la exposición a los sesgos cognitivos. Así, Shenhav et al. (2017) hallaron que el TRC es un predictor válido de la preferencia por una gratificación monetaria inmediata frente a una gratificación mayor pospuesta en el tiempo. Sin embargo, los resultados reflejaron que la capacidad cognitiva no se asoció con el descuento temporal. Los resultados indican que la capacidad cognitiva no puede explicar la correlación entre el TRC y el descuento temporal y que, por ende, el TRC no es una medida exclusiva de capacidad cognitiva.

Tabla 5.

Sesgos Cognitivos, Definición y Relación con el TRC

Sesgos cognitivos	Definición	Relación con RC	Tamaños del efecto	Referencias
Preferencia temporal (<i>Time preferences</i>)	Hace referencia a las preferencias de los individuos a recibir un pago inmediato o un pago superior en un futuro. Relacionado con los estudios del retraso o aplazamiento de las gratificaciones de Mischel (1996).	Puntuación más alta en el TRC: más paciencia en el retraso de las gratificaciones.	Rango: .10 a .29	Frederick (2005); Kuhn et al. (2017); Oechssler et al. (2009); Reyna y Wilhelms (2017); Shenhav et al. (2017).
Actitud hacia el riesgo: Aversión al riesgo o Búsqueda de Riesgo (<i>Risk aversion or risk seeking</i>)	Tendencia a percibir una pérdida monetaria más aversiva de lo que puede resultar atractiva una ganancia de la misma cantidad monetarias. Por lo tanto, es la tendencia, a aceptar una ganancia segura frente a una probabilidad media-alta de ganar una cuantía de mayor valor (dominio de las ganancias) o la tendencia a arriesgar en la toma de decisión bajo incertidumbre cuando existe una posibilidad de ganar, a pesar de que la probabilidad de perder es elevada (dominio de las pérdidas).	Dominio de las ganancias. Puntuación más alta en el TRC: menos aversión al riesgo. Dominio de las pérdidas: Resultados no estadísticamente significativos.	Rango: .05 a .26 Rango: .04 a .17	Campitelli y Labollita (2010); Cueva, et al. (2015); Ferrara et al. (2015); Frederick (2005); Instler, et al. (2013, 2015); Noori (2016); Thomson y Oppenheimer (2016).
Ambigüedad de riesgo	Es la preferencia por el riesgo sobre prospectivas ambiguas que son equivalentes desde un visto de vista de la utilidad esperada subjetiva.	Puntuación más alta en el TRC: menor preferencia por la ambigüedad de riesgo.	Rango: .19 y .23	Cesarini et al. (2012); Fairley y Weitzel (2017).
Necesidad de cierre cognitivo (<i>Need for Closure</i>)	Mide una motivación interna para evitar la ambigüedad de información y alcanzar una respuesta clara a las preguntas o situaciones.	Puntuación más alta en el TRC: menor aversión a la ambigüedad de la información	Rango: .16 a .22	Beatty y Thompson (2012); Koscielniak et al. (2016); Poore et al. (2014); Ventis (2015).

Continúa

Tabla 5.
Continuación

Sesgos cognitivos	Definición	Relación con RC	Tamaños del efecto	Referencias
Falacia de conjunción (<i>Conjunction fallacy</i>)	Tendencia a creer que es más probable que se dé una situación en la que dos o más eventos conjuntamente convergen (cajera de banco y activista) respecto a que se dé solamente uno de esos eventos (cajera de banco). Sin embargo, especificar un posible suceso con mayor detalle, solo puede reducir su probabilidad. Se da así, un conflicto entre la intuición de representatividad y la lógica de la probabilidad.	Puntuación más alta en el TRC: menor comisión de la falacia.	Rango: .10 a .16	Aczel et al. (2015); Alós-Ferrer et al. (2016); Alós-Ferrer y Hügelschäfer (2016); Noori (2016); Oechssler et al. (2009).
Sesgo de Conservadurismo (<i>Conservatism bias</i>)	Se da cuando una persona tiende a subestimar altas probabilidades y a sobreestimar bajas probabilidades	Puntuación más alta en el TRC: menor comisión del sesgo.	Ej. $r = .11$	Albaity, et al. (2014); Alós-Ferrer y Hügelschäfer (2016); Hoppe y Kusterer (2011); Noori (2016); Obrecht et al. (2009).
Efecto Anclaje (<i>Anchoring Effect</i>)	Efecto que se produce cuando las personas consideran un valor particular para una cantidad desconocida antes de estimar esa cantidad.	No asociación entre RC y este sesgo.		Bergman et al. (2010); Oechssler et al. (2009); Teovanović et al. (2015).
Efecto Enmarque (<i>Framing Effect</i>)	Se refiere a la tendencia de las personas a juzgar de manera diferente la información objetiva cuando está redactada de manera diferente	No consenso entre los resultados.	Rango: .00 a .27	Aczel et al. (2015); Jackson et al. (2016); Primi et al. (2015); Sinayev y Peters (2015); Toplak et al. (2014); Weller et al. (2013).
Insensibilidad al tamaño de la muestra o “Ley de los Pequeños Números”	Tendencia a no tener en cuenta que los resultados de muestras grandes merecen más confianza que los de muestras pequeñas, puesto que estas arrojan resultados extremos con más frecuencia que las muestras grandes.	Puntuación más alta en el TRC: menos insensibilidad al tamaño de la muestra.	Rango: .11 a .41	Aczel et al. (2015); Cesarini et al. (2012); Obrecht et al. (2009).

Continúa

Tabla 5.
Continuación

Sesgos cognitivos	Definición	Relación con RC	Tamaños del efecto	Referencias
Sesgo de <i>status quo</i> o de opciones por defecto (<i>Default Bias</i>)	Tendencia a aceptar las opciones predeterminadas, sin deliberar hacia otras posibilidades	Puntuación más alta en el TRC: menor exposición al sesgo.	Ej. $r = .08$	Altmann y Falk (2009); Borghans y Golsteyn (2014); Cesarini et al. (2012).
Probabilidad de unión (<i>Probability Match</i>)	Es una tarea de elección probabilística donde los participantes presentan una tendencia a escoger la opción de respuesta que ofrece una mayor ganancia, sin examinar la probabilidad de ocurrencia del evento (estrategia de unión), en lugar de seleccionar la opción de respuesta que ofrece la mayor probabilidad de que suceda el evento (estrategia maximizadora). Es decir, es la tendencia a unir la probabilidad de elección con la probabilidad del resultado.	Puntuación más alta en el TRC: mayor empleo de la estrategia maximizadora.	Rango: .23 a .40	Aczel et al. (2015); Kochler y James (2010); Newell et al. (2013); Shenhav et al. (2017).
Efecto “1-de-X”	Se puede definir como la tendencia de los sujetos a percibir más alta la probabilidad expresada en el formato “1-de-X” (por ejemplo, 1 de 10) que en el formato “1-de-X*N” (por ejemplo 10 de 100) a pesar de tratarse de la misma probabilidad.	No asociación entre RC y este efecto.		Sirota et al. (2014).
Sesgo de ignorar tasa base o ignorar proporciones (<i>Base Rate Neglect</i>)	Tendencia a ignorar la información estadística y probabilística, dándole mayor peso a la evidencia de un caso individual específico. Envuelve un conflicto entre estereotipos intuitivos (heurística de la representatividad) y el ratio de probabilidad.	Puntuación más alta en el TRC: menos susceptibles al sesgo de tasa base.	Rango: .11 a .50	Cesarini et al. (2012); Hoppe y Kusterer (2011); Pennycook et al. (2014a); Teovanović et al. (2015); Thompson et al. (2018); Weller et al. (2013).
Sesgo de Ignorar el Denominador (<i>Denominator Neglect Bias</i>)	Consiste en escoger la frecuencia absoluta más alta incluso cuando no se corresponde con la proporción más alta	Puntuación más alta en el TRC: menor tendencia a ignorar el denominador.	Rango: .37 a .42	Thomson y Oppenheimer (2016); Toplak et al. (2014).

Continúa

Tabla 5.
Continuación

Sesgos cognitivos	Definición	Relación con RC	Tamaños del efecto	Referencias
Sesgo de Proporción Dominante (<i>Proportion Dominance</i> , <i>Reference Group Effect</i> , <i>Drop-in-the-Bucket Thinking</i> o <i>Psychophysical Numbing</i>)	Preferencia por la opción de salvar 225 sujetos de 300 en lugar de salvar 230 de 920. Es decir, se da más peso a la ganancia de la proporción (escogemos la proporción más alta en sí) que a la ganancia del resultado absoluto (ignoramos la frecuencia absoluta). Es el sesgo opuesto al sesgo de Ignorar el Denominador.	Puntuación más alta en el TRC: menor tendencia a cometer el sesgo.	Ej. $r = .26$	Mata (2016).
Sesgo de sobre confianza/exceso de seguridad (<i>Overconfidence Bias</i>)	Es la común inclinación por sobrestimar nuestras propias habilidades en el éxito del desempeño de una particular tarea (una excesiva auto seguridad en la decisión tomada).	No asociación o asociación prácticamente nula.	Ej. $r = .08$	Hoppe y Kusterer (2011); Langnickel y Zeisberger (2016); Noori (2016); Sinayev y Peters (2015); Teovanović et al. (2015).
Sesgo del Punto Ciego (<i>Bias Blind Spot</i>)	Tendencia a pensar que, la comisión de sesgos cognitivos y motivacionales es más prevalentes en otros que en sí mismo.	Puntuación más alta en el TRC: mayor exposición al sesgo.	Rango: .096 y .102	Mata, Fiedler et al. (2013); Toplak et al. (2014). West et al. (2012).
Efecto Asignación (<i>Endowment Effect</i>)	Es el hecho de que las personas a menudo exigen mucho más para renunciar a un objeto de lo que estarían dispuestos a pagar para adquirirlo.	No asociación entre TRC y el efecto asignación.		Hoppe y Kusterer (2011).
Error del Coste invertido o perdido (<i>Sunk Cost Effect</i>)	Tendencia a continuar una tarea una vez que se haya hecho una inversión de dinero, esfuerzo o tiempo. Consiste en describir varias situaciones en las cuales la persona que debe tomar la decisión tiene que escoger entre dos opciones: una en la que no es posible recuperar un gasto pasado (<i>sunk-cost option</i>) y la otra que es más beneficiosa en términos económicos (<i>normative option</i>).	Puntuación más alta en el TRC: menor susceptibilidad al sesgo	Rango: .10 y .22	Aczel et al. (2015); Cesarino et al. (2012); Teovanović et al. (2015).

Continúa

Tabla 5.
Continuación

Sesgos cognitivos	Definición	Relación con RC	Tamaños del efecto	Referencias
Efecto <i>Knobe</i>	Hace referencia a la tarea “daño” (<i>harm</i>) donde se propone un dilema moral sobre si aceptar un programa que facilitaría la vida de personas en silla de ruedas, pero dañaría el medio ambiente.	Puntuación más alta en el TRC: menor aceptación del programa (menor efecto <i>Knobe</i>).	Ej. $r = .148$	Pinillos et al. (2011).
Sesgo de Creencia (<i>Belief/Bias</i>)	Previsible tendencia a evaluar argumentos deductivos sobre la base de la credibilidad de la conclusión por encima de la base de su validez lógica.	Puntuación más alta en el TRC: menor comisión del sesgo de creencia en el razonamiento silogístico.	Rango: .17 a .68	Eriksson y Jansson (2016); Hopkins et al. (2016); Howarth et al. (2016); Szasz et al. (2017); Teovanović et al. (2015); Thomson y Oppenheimer (2016); Toplak et al. (2014, 2017); Trémolière et al. (2014); Viator et al. (2014).
Sesgo de referencia o estatus social	Hace referencia a la tendencia de los sujetos a dejarse influenciar por el estatus social o el estatus referente de las personas involucradas en la toma de decisión. Así, por ejemplo, existe mayor tendencia a aceptar una invitación procedente de un famoso de televisión que de un vecino desconocido.	Puntuación más alta en el TRC: menor exposición al sesgo	Ej. $r = .44$	Baldi et al. (2013).
Ilusión de Control (<i>Illusion of control</i>)	Valoración de si las apuestas sobre resultados aleatorios están afectadas por la cantidad de control que los encuestados perciben que ejercen sobre el resultado de una lotería.	Puntuación más alta en el TRC: menor percepción de control	Rango: .17 y .24	Cesarini et al. (2012); Noori (2016); van Prooijen (2017).
Falacia del Jugador (<i>Gambler's Fallacy</i>)	Se refiere a la tendencia a creer que, si algo sucede con una frecuencia más alta de lo normal, en ocasiones posteriores debe suceder con menos frecuencia	Puntuación más alta en el TRC: menor comisión de la falacia.	Rango: .11 y .26	Aezel et al. (2015).

Continúa

Tabla 5.
Continuación

Sesgos cognitivos	Definición	Relación con RC	Tamaños del efecto	Referencias
Sesgo de Retrospección (<i>Hindsight Bias</i>)	Es la tendencia a percibir eventos como más probables una vez que han sucedido. Es decir, los juicios hechos a partir de información acerca del resultado de un evento difieren de los juicios hechos sobre ese mismo evento sin dicha información.	No asociación entre las puntuaciones en el TRC y la exposición al sesgo.	Ej. $r = .06$	Teovanović et al. (2015).
Sesgo de Resultados (<i>Outcome Bias</i>)	Tendencia a juzgar la calidad de las decisiones basadas en la información sobre los resultados de dichas decisiones.	Puntuación más alta en el TRC: menor exposición al sesgo	Rango: .12 y .23	Aczel et al. (2015); Teovanović et al. (2015).
Regresión a la media (<i>Regression to the Mean</i>)	Fenómeno estadístico que se refiere a que, tras un acontecimiento extremo, la siguiente medición más probable será un valor cercano a la media. También hace referencia a que, ante circunstancias de incertidumbre, la puntuación media es siempre la mejor predicción.	No asociación con las puntuaciones en el TRC		Aczel et al. (2015).

Primi et al. (2015) encontraron que la RC fue un predictor válido y significativo de la búsqueda de riesgo, incluso mejor que capacidad numérica e inteligencia. Sin embargo, Agnew y Harrison (2017) encontraron que, una vez controlado por las variables demográficas, los factores de personalidad y la capacidad cognitiva, el TRC no tuvo efectos significativos en la predicción de la disposición hacia el riesgo. Asimismo, el análisis de regresión llevado a cabo por Fairley y Weitzel (2017) para predecir la ambigüedad de riesgo demostró que el TRC es un predictor independiente, incluso una vez controlado por otras variables como, por ejemplo, aversión al riesgo, conocimiento financiero y variables demográficas.

En el estudio de Kochler y James (2010) se encontró que los sujetos más reflexivos mostraron una mayor probabilidad a escoger la estrategia maximizadora en las tareas de probabilidad de unión que los sujetos más intuitivos. Esta relación se mantuvo estadísticamente significativa una vez controlado los resultados por los efectos de la capacidad matemática. Los autores defienden que el TRC es una medida de capacidad cognitiva pero también de disposición de pensamiento. De modo que la intuición inicial de unir la probabilidad de elección con la probabilidad del resultado debe ser inhibida a favor de la opción que aporte un resultado de mayor probabilidad, mostrando, así, una mayor tendencia al análisis de la respuesta inicial (disposición de pensamiento) y siendo capaz de generar una respuesta estratégica alternativa superior (capacidad cognitiva).

Del mismo modo, Pennycook et al. (2014a), al examinar la relación entre el TRC y la tarea del sesgo de tasa base incongruente⁵, encontraron que el TRC correlacionó positiva y significativamente con la diferencia de tiempo invertida en responder al problema de tasa base incongruente respecto del problema de tasa base congruente. Los resultados reflejaron que los sujetos con puntuaciones más altas en el TRC son más propensos a detectar y resolver correctamente el conflicto entre creencias estereotipadas y tasas base. Sin embargo, la medida de capacidad cognitiva verbal no se relacionó con el tiempo invertido en la respuesta al problema, sugiriendo que la disposición de pensamiento, pero no la capacidad cognitiva predice la detección de conflicto en los problemas de razonamiento.

Toplak et al. (2011) hallaron que el TRC explicó varianza única de un compuesto de sesgos cognitivos una vez controlado por los efectos de IQ. Esto es, el TRC explicó varianza sustancial del pensamiento racional que no pudo ser explicada por otras medidas de capacidad cognitiva o funcionamiento ejecutivo. No obstante, Szasz et al. (2017) encontraron que, después de controlar por los efectos de la capacidad numérica, la comisión del sesgo de creencia silogística no se asoció con el desempeño en el TRC.

⁵ Conjunto de problemas con información estereotipada que entra en conflicto con la información probabilística o con la información de la tasa base ofrecida por el problema.

Asimismo, Sinayev y Peters (2015) aplicaron el modelo matemático de miseria cognitiva de Böckenholt (2012) denominado Modelo de Respuesta Cognitiva-Avara (*Cognitive-Miser Response Model*). Este modelo asume que el TRC es una medida de RC y de capacidad numérica. La medida de RC (capacidad meta-cognitiva de supresión de la respuesta inmediata) está compuesta por la suma de respuestas no intuitivas, es decir, respuestas correctas y erróneas no intuitivas; y la medida de capacidad matemática (capacidad cognitiva) está compuesta por el número total de respuestas correctas. En base a ello, encontraron correlaciones ligeramente diferentes entre esas dos medidas y algunas tareas de sesgos cognitivos. La medida de cálculo mostró correlaciones ligeramente superiores con las tareas de sesgos cognitivos que la medida de RC. Adicionalmente, los análisis de regresión mostraron que, en todos los casos, la medida de cálculo fue un predictor válido y fiable de la exposición a los sesgos, pero no la medida de RC. Asimismo, el análisis de regresión del estudio 2 mostró que, una vez controlado por los efectos de las variables demográficas y la capacidad mental general, la medida de cálculo y *numeracy* predijeron los sesgos de decisión de forma independiente. Sin embargo, la medida de RC no tuvo efectos significativos. Los resultados obtenidos en este estudio sugieren que el TRC es, en primer lugar, una medida de capacidad numérica.

Para resumir, los resultados expuestos anteriormente han evidenciado que hay varianza de pensamiento racional que es explicada por la RC independientemente de la capacidad cognitiva y, por ende, el TRC puede ser un test que mide propiedades relevantes de pensamiento analítico más allá de los aspectos que miden los test de inteligencia. Toplak et al. (2011) indican que la mente humana tiene dos características que nos hacen menos racionales: (1) el problema de procesamiento y (2) el problema de contenido. El problema de procesamiento hace referencia a que los humanos tendemos a ser tacaños cognitivos porque nuestra tendencia básica es a procesar por defecto con mecanismos heurísticos o que impliquen menor esfuerzo cognitivo. Consecuentemente, es necesaria la disposición a aplicar un pensamiento analítico. El problema de contenido hace referencia a la necesidad de poseer el conocimiento declarativo y las reglas estratégicas necesarias para inhibir la respuesta intuitiva y reemplazarla por otra más precisa en términos cognitivos. Por lo tanto, el proceso de inhibición no es solo procedimental sino también de contenido. Tener lagunas en las estructuras de conocimiento implica un segundo tipo de error de razonamiento. Esta puede ser la explicación de por qué el TRC es en parte independiente de la capacidad cognitiva, ya que las medidas de CMG no evalúan la tendencia hacia un procesamiento cognitivo avaro.

2.3. ¿Es el TRC una Medida de Capacidad Cognitiva, Pensamiento Racional y Disposición hacia un Pensamiento de Mente Abierta?

Hasta el momento, se han presentado las dos principales hipótesis de qué mide el TRC. Las evidencias científicas respaldan, por un lado, que el TRC mide capacidad cognitiva y,

específicamente, capacidad numérica; lo que implica que el TRC no es sustancialmente diferente de otras medidas matemáticas y, por ende, la capacidad numérica conduce la capacidad del test para predecir mejores decisiones (Weller et al. 2013). Por otro lado, el TRC es identificado como una medida de RC tendente a comprobar y detectar errores intuitivos e inhibir la impulsiva respuesta inmediata de nuestro sistema 1 a favor de un procesamiento analítico. Recientemente, varios investigadores han tratado de demostrar que ambas propuestas no son excluyentes y que el proceso que subyace en el desempeño en el TRC podría incluir la RC y la capacidad numérica (Sinayev y Peters, 2015).

Cokely and Kelley (2009) asociaron el TRC con la reflexividad (*reflectiveness*), es decir, con un pensamiento o procesamiento minucioso, cuidadoso y con una cognición elaborada. Así, defienden que la aplicación de los mecanismos de control cognitivo podría jugar un importante rol en la reflexividad y en el mejor desempeño de tareas, a través de una cognición adaptativa. Campitelli y Labollita (2010) propusieron que, la RC no es solo una capacidad o disposición para anular la respuesta dominante, sino que, también se refiere a la capacidad y disposición a iniciar el procesamiento cognitivo. Además, propusieron que la RC implica un pensamiento más cuidadoso, minucioso y elaborado y menos impulsivo. Está, por tanto, relacionada con el amplio concepto de *Actively Open-Minded Thinking* (AOT) de Baron (1988). El concepto de AOT hace referencia a la disposición a buscar más posibilidades antes de hacer una inferencia. Haran, Ritov y Mellers (2013) lo definen como un “estilo cognitivo que mide la tendencia a sopesar nuevas evidencias frente a las creencias del sujeto, a emplear el tiempo suficiente en el problema antes de rendirse y a considerar detenidamente las opiniones de otros en la formación de las de uno mismo” (pág. 189). Así, Campitelli y Labollita (2010) sugieren que el desempeño en el TRC recae en el pensamiento activo de mente abierta o en la búsqueda de respuestas alternativas. En cualquier caso, ambas explicaciones indican que el éxito en el desempeño en el TRC se apoya en un procesamiento analítico adicional que pueda debilitar la inadecuada respuesta intuitiva (Pennycook, Cheyne, Koehler et al., 2015) y defienden que el TRC no solo mide capacidad sino también una disposición de pensamiento.

Juanchich et al. (2016) van un paso más allá e indican que el TRC también mide un amplio conjunto de disposiciones. Los autores defienden la suposición de que el TRC evalúa la tendencia a ser exhaustivo y/o a participar activamente en el pensamiento de mente abierta. Pero, además, afirman que el TRC mide la capacidad más genérica de usar la intuición cuando esta sea apropiada. En este contexto, la RC es, en psicología, un elemento para seleccionar el proceso cognitivo más adecuado a la situación específica. Por lo tanto, de acuerdo con esta suposición, el TRC pertenecería a la clase de pruebas que miden reflexión-impulsividad en su sentido más general (Baron et al., 2015).

Macchi y Bagassi (2012) también apoyan esta hipótesis y afirman que las diferencias en el desempeño en el ítem del bate y la pelota son debidas a un tipo muy concreto de reflexibilidad

cognitiva: la competencia pragmática, ya que la tarea se interpreta a la luz del contexto de la presentación de la información en el enunciado del problema y no se abstrae o se descontextualiza de ella. Por lo que, una mente abierta podría encontrar más fácilmente la correcta solución a los problemas, al enfocarlos desde otra perspectiva.

Consecuentemente, varios estudios han explorado la relación entre la RC y la medida de AOT. Haran et al. (2013) hallaron una asociación de .30 ($N = 183$, $p < .001$). Thompson et al. (2013) encontraron una correlación de .32 ($N = 142$, $p < .001$). Toplak et al. (2014) mostraron que tanto el TRC original como el desarrollado por los autores se asociaron positiva y significativamente con AOT ($r(160) = .29$ y $r(160) = .42$, respectivamente). Asimismo, otros investigadores también encontraron correlaciones moderadas con AOT (Baron et al., 2015; Svedholm-Häkkinen y Lindeman, 2018; Toplak et al., 2017).

Las puntuaciones del TRC también se han asociado con otras medidas indicadoras de pensamiento de mente abierta. Por ejemplo, Barr et al. (2015) examinaron la relación entre el TRC y el pensamiento creativo, medido por diversas tareas que reflejan las multi facetas del constructo de creatividad: flexibilidad cognitiva, fluidez y originalidad y conexiones remotas. Los resultados mostraron que el TRC fue un predictor válido de todas las medidas de creatividad. El análisis de regresión para predecir el compuesto de creatividad mostró que el compuesto de estilo cognitivo analítico (TRC y tareas de tasa base) y la capacidad cognitiva fueron predictores válidos e independientes de la medida de creatividad.

Asimismo, Corgnet, Espín y Hernán-González (2015) encontraron que el TRC correlacionó significativamente con la medida de pensamiento creativo convergente ($r(150) = .17$) pero no con las medidas de pensamiento creativo divergente. Los análisis de regresión lineal mostraron que, cuando se controló por los efectos de la inteligencia, el TRC no explicó varianza del pensamiento creativo convergente más allá de la explicada por CMG, lo que indica que toda la varianza que el TRC determina es debido a la varianza compartida con inteligencia. Sin embargo, el análisis de regresión cuadrático mostró que el TRC sí predice significativamente algunas medidas de pensamiento creativo divergente. Estos datos sugieren que los sujetos con puntuaciones promedio en el TRC tienden a producir más respuestas y a usar más categorías que aquellos sujetos con alta o baja RC. Controlando los análisis por los efectos de la inteligencia, esta relación no se vio alterada y se mantuvo significativa, por lo que el estudio replicó parcialmente los resultados de Barr et al. (2015).

Gino y Ariely (2012) investigaron la relación entre el TRC y el estilo de pensamiento creativo, utilizando 3 medidas de pensamiento creativo: la escala de personalidad creativa, el inventario de comportamiento creativo y una escala de estilo cognitivo creativo. Los resultados indicaron que el TRC no correlacionó significativamente con ninguna de esas medidas.

Del mismo modo, Del Missier, Visentini y Mantyla (2015) no encontraron una asociación entre las puntuaciones en el TRC desarrollado por Finucane y Gullion (2010) y tres medidas de

pensamiento abierto relacionadas la toma de decisión en escenarios reales. Los resultados mostraron que el TRC no correlacionó de forma significativa con las medidas de fluidez de generación de opciones ni con las medidas de diversidad. Sin embargo, se asoció significativamente con la aplicación de reglas de decisión (ARD; *Applying Decision Rules*). Los resultados sugirieron que el ARD y la generación de ideas subyacen de procesos cognitivos diferentes. Mientras que el primero está más basado en el control, la memoria y el razonamiento analítico, el segundo depende en mayor medida de las habilidades de creación de ideas. Por lo tanto, ambos parecen representar diferentes facetas de la competencia de toma de decisión.

Por otro lado, y partiendo de la suposición de que el TRC mide la capacidad más genérica de seleccionar y aplicar el proceso cognitivo o la disposición de pensamiento más adecuado a cada situación (Juanchich et al., 2016), diversos estudios han examinado la relación entre el TRC y otros estilos cognitivos y de toma de decisión.

La medida de estilo cognitivo que ha sido más examinada en relación con el TRC ha sido la necesidad de cognición, medida con la escala *Need for Cognition* (NFC; Cacioppo y Petty, 1982; Cohen, Scotland y Wolfe, 1955). NFC fue definida por Cohen y colaboradores (1955) como “la necesidad de entender y hacer razonable la experiencia con el mundo” (pág. 291). En un momento posterior, Cacioppo y Petty (1982) adoptaron este término y propusieron que la necesidad de cognición hace referencia a la “tendencia a disfrutar de las actividades que implican esfuerzo cognitivo” (pág. 116). En esta misma línea, Haran et al. (2013) definieron la NFC como un estilo cognitivo que se refiere a la tendencia a comprometerse y disfrutar con el esfuerzo cognitivo.

Frederick (2005) encontró que el TRC correlacionó positiva y significativamente con NFC. Además, halló que los sujetos con altas puntuaciones en NFC correlacionaron mejor con problemas de aritmética, anagramas, trabajos escolares, aprenden más, recuerdan mejor la información a la que están expuestos, generan más conocimientos relevantes para la tarea y disfrutaban en mayor medida en el procesamiento de la información que los sujetos con bajas puntuaciones en la escala. También encontró que NFC se asoció de forma más robusta con el TRC que con las medidas de CMG. Estudios posteriores han obtenido resultados similares a los encontrados por Frederick (Barr et al., 2015a; Haran et al., 2013; Iyer, Koleva, Graham, Ditto y Haidt, 2012; Lindeman y Svedholm-Häkkinen, 2016; Primi et al., 2017; Teovanović et al., 2015; Thomson y Oppenheimer, 2016; Toplak et al., 2014; Ventis, 2015; Viator et al., 2014).

Posteriormente, Pacini y Epstein (1999) ampliaron en 20 ítems la escala de NFC y la integraron como parte del inventario racional-experiencial (*Rational-Experiential Inventory*, REI), unido a la escala *Faith in Intuition* (FI). La escala de FI fue originalmente desarrollada por Epstein, Pacini, Denes-Raj y Heier (1996), está basada en los MPD de psicología y economía y refleja la confianza de los sujetos en su propia intuición. La escala de FI es una medida de autoinforme que refleja el compromiso y la confianza en nuestra propia intuición y, por tanto, es

una medida de estilo cognitivo subjetiva en oposición a una medida objetiva como es el TRC (Alós-Ferrer y Hügelschäfer, 2016).

En relación con ello, Liberali et al. (2012) encontraron una asociación entre el TRC y la escala de NFC de REI pero no con la escala de FI. Calvillo y Burgeno (2015) hallaron que el TRC de 7 ítems desarrollado por Toplak et al. (2014) correlacionó de forma significativa con las subescalas de REI racional pero no correlacionó con las subescalas de razonamiento experiencial.

Del mismo modo, otros estudios que han examinado la relación entre la RC y la medida de REI han reflejado una tendencia a que las puntuaciones del TRC se asocien con la medida racional pero no con la medida experiencial (Alós-Ferrer y Hügelschäfer, 2016; Finucane y Gullion, 2010; Poore et al., 2014; Razmyr y Reeve, 2013; Szaszi et al., 2017; Welsh et al., 2013). Solamente en dos casos, se ha encontrado una relación negativa y significativa entre las puntuaciones del TRC y la escala FI (Alós-Ferrer y Hügelschäfer, 2016; Lindeman y Svedholm-Häkkinen; 2016).

También se ha encontrado que FI es independiente del compromiso o disfrute con la actividad cognitiva, medido a través de la escala NFC y que no produce efectos significativos en la predicción de los sesgos cognitivos (Alós-Ferrer y Hügelschäfer, 2016; Pennycook, Cheyne, Koehler et al., 2015), quizás por ello, el TRC no tiende a asociarse con la escala de FI. Las evidencias empíricas sugieren que el TRC no debe ser tomado como una medida exactamente opuesta a la escala de FI ya que los datos indican una nula asociación entre ambas medidas (Alós-Ferrer y Hügelschäfer, 2012).

Sin embargo, diversos investigadores han comenzado a utilizar el TRC como un instrumento que puede medir la RC y la impulsividad (o un *proxy* del constructo), donde aquellos que no responden ningún ítem del TRC correctamente son altos en impulsividad comparados con los sujetos que responden correctamente las tres preguntas del TRC (Agnew y Harrison, 2017). El sumatorio de ítems respondidos incorrectamente se emplea como un indicador de impulsividad (Brosnan et al., 2014; Piazza y Sousa, 2014; Shenhav et al., 2011). Bajo esta premisa, el TRC se configura como una medida de racionalidad-impulsividad, en su sentido más amplio, empleando el número de ítems respondidos correctamente o el número de errores cometidos, respectivamente. No obstante, Pennycook, Cheyne, Koehler et al. (2015) consideran que esta última estrategia no es correcta, ya que basándose en la relación existente entre el TRC y las medidas de autoinforme de estilos cognitivos analíticos-impulsivos (tales como NFC o FI), se encontró que el TRC, medido como reflexividad (número de ítems respondidos correctamente), se asoció con estas medidas. Sin embargo, el TRC medido como impulsividad (número de ítems respondidos incorrectamente) no correlacionó con ellas. Los autores validaron el TRC como medida de impulsividad y encontraron que tanto el TRC reflexivo como el TRC impulsivo correlacionaron significativamente con NFC y con FI y las correlaciones entre el TRC reflexivo y el TRC impulsivo con las medidas de REI fueron prácticamente iguales. No obstante, la relación

de FI debería ser más robusta con el TRC intuitivo que con el TRC reflexivo; y este último debería mostrar una mayor correlación con NFC que el TRC impulsivo, al tratarse de constructos más homogéneos. Del mismo modo, los autores encontraron que la proporción de respuestas intuitivas incorrectas correlacionó significativamente con el TRC intuitivo, pero no con el TRC reflexivo. Igualmente, FI no fue más alto para los sujetos que dieron una respuesta intuitiva incorrecta que para los sujetos que proporcionaron una respuesta incorrecta diferente de la intuitiva. Sin embargo, los sujetos más intuitivos deberían responder en mayor medida con la respuesta intuitiva incorrecta que con otro tipo de respuesta incorrecta. Esto hace cuestionarse la validez de la media intuitiva del TRC como una medida de impulsividad. Los autores defienden que el TRC no es un test de intuición sino de reflexibilidad o disposición hacia un razonamiento analítico. En cualquier caso, sugieren que el TRC podría indicar preferencia por lo intuitivo, pero no medir el rasgo de impulsividad en sí mismo. Así, sostienen que, los sujetos intuitivos podrían, o no, detectar la necesidad de pensar analíticamente sobre el problema, pero decidir, sin embargo, actuar de acuerdo con sus intuiciones.

Finalmente, también se ha examinado la relación entre el TRC y otras disposiciones de pensamiento o de toma de decisión. Por ejemplo, Juanchich et al. (2016) examinaron la relación entre las puntuaciones en el TRC y el cuestionario de estilos de decisión de Leykin y DeRubeis (2010; *Decision Styles Questionnaire*) y encontraron que, de los 9 estilos de decisión, el TRC solo correlacionó significativamente con los estilos de toma de decisión intuitivo y espontáneo. Por su parte, Toplak et al. (2017) hallaron que el TRC correlacionó significativamente con pensamiento orientado al futuro y con la evitación a un pensamiento supersticioso. Toplak y colaboradores (2014) encontraron resultados en esta misma línea. El estudio de van Prooijen (2017) mostró que el TRC correlacionó significativamente con el pensamiento hacia la creencia en soluciones simples y la creencia en teorías conspiratorias. Finalmente, testaron si la creencia en soluciones sencillas es atribuible al pensamiento analítico (RC). El análisis de mediación demostró que la educación decrece las creencias en soluciones sencillas debido a un incremento de la capacidad de pensamiento analítico.

En conjunto, las evidencias empíricas sugieren que el TRC es una tarea compleja, que recoge propiedades de personalidad, en cuanto a las disposiciones de pensamiento o estilos cognitivos (bien sea de mente abierta o de tendencia hacia un pensamiento racional), y propiedades de capacidad cognitiva (como podría ser la capacidad matemática, la memoria operativa y las funciones metacognitivas de monitorización e inhibición). De modo que la potencia predictiva del TRC podría derivar de ambas dimensiones ya que, se ha demostrado que el TRC se relaciona con medidas de CMG y también implica disposiciones de pensamiento, especialmente aquellas relacionadas con la reflexividad, la tendencia a disfrutar de tareas cognitivas complejas y la tendencia a buscar soluciones alternativas (Toplak et al. 2011).

En este sentido, diversos autores han examinado cómo contribuyen diferentes variables en la determinación de la RC. Finucane y Gullion (2010) llevaron a cabo una regresión jerárquica para predecir el desempeño en el TRC y encontraron cambios significativos en R-cuadrado al ir introduciendo las variables demográficas, diversas medidas de salud y las variables de capacidad cognitiva. No obstante, las medidas de estilos cognitivos no incrementaron significativamente la R-cuadrada.

Toplak et al. (2014) llevaron a cabo un análisis de regresión para examinar si la CMG y el compuesto de las disposiciones de pensamiento determinan la RC. Los resultados mostraron que, en conjunto, ambas medidas explicaron el 31% de la varianza de las puntuaciones en el TRC. La CMG fue el mejor predictor, pero ambas variables explicaron varianza única del test.

Asimismo, Szasz et al. (2017) encontraron que, después de controlar por la capacidad numérica, el análisis de correlaciones parciales mostró que solo REI y AOT mantuvieron una relación significativa con el TRC. Sin embargo, encontraron que el modelo final con mejor ajuste se compuso por la capacidad numérica y por la medida de REI. Por lo tanto, los resultados de este estudio apoyan la idea de que la definición de RC, operacionalizada mediante el TRC, no debería ser restringida a la descripción de capacidad o disposición para inhibir las intuiciones, sino también como una medida amplia de reflexividad-impulsividad.

Por otro lado, Welsh et al. (2013) llevaron a cabo un análisis factorial, donde los resultados reflejaron que el TRC cargó hacia el mismo factor que las medidas de capacidad cognitiva (capacidad numérica, velocidad de procesamiento y memoria operativa). Asimismo, el TRC también cargó, aunque en menor medida, hacia el factor compuesto por las medidas de estilo cognitivo o de toma de decisión (apertura a la experiencia (AP), *Need for Cognition* (NFC), Compromiso racional (REI), Racionalidad y capacidad racional). No obstante, la medida de metacognición sobre funcionamiento ejecutivo no cargó hacia ningún factor en el que se encontrase el TRC, lo que sugiere que no se asocia con esta medida. De acuerdo con los resultados, los autores concluyeron que, el TRC parece ser una medida de capacidad que permite reconocer los errores cometidos en las respuestas intuitivas numéricas, más que una función de metacognición. El estudio muestra evidencias de que el TRC mide capacidad cognitiva numérica ya que, fue con esta variable con la que se asoció de forma más robusta. Del mismo modo, demostró ser una medida de estilo cognitivo relacionada con la disposición hacia un pensamiento analítico, pero no mostró su valencia como medida de pensamiento impulsivo.

Por último, en el análisis factorial llevado a cabo por Teovanović y colaboradores (2015) con medidas de capacidad cognitiva y disposiciones de pensamiento, se encontró que el TRC no cargó hacia ninguno de los tres factores obtenidos. Las cargas factoriales del TRC no llegaron a .40 en ninguno de ellos. No obstante, residualmente cargó hacia el factor de IQ fluida, mientras que no cargó hacia el factor de disposiciones de pensamiento. Simultáneamente, se llevó a cabo un nuevo análisis factorial añadiendo, a las anteriores, las medidas de sesgos cognitivos. Los

resultados reflejaron que, en este caso, el TRC cargó de forma equitativa entre el factor de IQ fluida y los dos nuevos factores sobre sesgos cognitivos. También cargó hacia el factor de disposición de pensamiento, aunque en menor medida. Estos datos sugieren una doble dimensionalidad del TRC: la capacidad cognitiva y la tendencia a inhibir la respuesta heurística y, marginalmente, se sugirió la dimensionalidad de disposición de pensamiento.

En resumen, los resultados anteriores han mostrado que existe una moderada relación entre el TRC y las medidas de capacidad cognitiva, especialmente con capacidad matemática. No obstante, los estudios también han indicado que, en numerosas ocasiones y para diferentes comportamientos y situaciones, la capacidad matemática no es suficiente para explicar las evidencias de validez del TRC. Frederick (2005) encontró que el TRC es superior a las medidas de capacidad mental general en la predicción de la comisión de errores y sesgos cognitivos debido al inherente procesamiento cognitivo. De este modo, el TRC se configura como una medida de disposición hacia un pensamiento analítico, donde el sujeto debe inhibir la respuesta inmediata e incorrecta de nuestro sistema 1 para dar paso a un procesamiento deliberado y racional del sistema 2. Diversos autores identifican al TRC con las funciones metacognitivas de monitorización e inhibición al exigir este un control y auto regulación del proceso cognitivo en sí mismo. Consecuentemente, el test se propuso como el mejor predictor de los juicios y toma de decisión, incluso superiores a los test de inteligencia. Este descubrimiento ha permitido sugerir que el estilo de decisión o la preferencia de los sujetos por un pensamiento racional o intuitivo podría ser más importante que la inteligencia al predecir la exposición a los sesgos cognitivos (Welsh et al., 2013). Finalmente, algunas evidencias han señalado que el TRC puede capturar ciertas características de un pensamiento de mente abierta, donde enfocar los problemas desde distintas perspectivas podría ayudar a adaptar el estilo de decisión más apropiado a cada situación y alcanzar mejores resultados.

3. EVIDENCIAS DE VALIDEZ BASADAS EN LA RELACIÓN CON OTRAS VARIABLES

3.1. Personalidad

3.1.1. Modelo de los Big Five Factors (BFF)

La relación del TRC con los cinco factores de personalidad no ha sido muy investigada y, aunque los resultados son poco ilustrativos, parecen indicar que el constructo de RC es independiente de los factores de personalidad. Uno de los factores de personalidad más estudiados en relación con la RC ha sido la Apertura a la Experiencia (AE; *Openness to Experience*), ya que evidencias científicas han demostrado que es el único factor de personalidad asociado con la capacidad cognitiva (Ackerman y Heggestad, 1997; Ashton, Lee, Vernon y Jang, 2000; Bates y Shieles, 2003; Wolf y Ackerman, 2005). La AE es definida como el grado en que un sujeto es

abierto a nuevas ideas, enfoques y experiencias y está asociada con la creatividad, visiones sociales liberales, fantasías y la necesidad de ampliar y examinar experiencias (Barrick y Mount, 1991; Costa y McCrae, 1997; Digman, 1990; Goldberg, 1992; Hogan, 1991). De acuerdo con la definición de AE este factor encajaría en la suposición de que el TRC no mide únicamente capacidad mental general y pensamiento racional, sino que también es un indicador de un pensamiento de mente abierta (AOT). No obstante, los resultados muestran correlaciones muy pequeñas o nulas entre el TRC y ese factor, sugiriendo que, aunque la activación mental (un estilo cognitivo) y la apertura a la experiencia (un rasgo de personalidad) comparten similitudes superficiales, son rasgos estadísticamente independientes (Browne, Pennycook, Goodwin y McHenry, 2014).

Del mismo modo, el TRC tiende a mostrar correlaciones pequeñas o nulas con el resto de los BFF, lo que sugiere que la RC y la personalidad son constructos independientes. Así, por ejemplo, Shenhav et al. (2011) no encontraron una relación significativa entre la RC y los factores de personalidad de AE y Extraversión (EX; *Extraversion*). Del mismo modo, examinaron la relación entre las puntuaciones intuitivas del TRC (número de ítems respondidos intuitivamente) y esos mismos factores de personalidad. Los resultados también reflejaron que el TRC intuitivo no correlacionó significativamente con ambos factores. Los resultados también indicaron que la relación con el TRC es ligeramente superior cuando se emplean las puntuaciones correctas que cuando se utilizan las puntuaciones intuitivas. Por su parte, Browne et al. (2014) encontraron que las puntuaciones en el TRC y la dimensión de AE no correlacionaron significativamente. Poore et al. (2014) hallaron una relación significativa con el factor de Amigabilidad (A, *Agreeableness*) y con el factor de conciencia (C, *Conscientiousness*). Sin embargo, las correlaciones con los restantes factores de personalidad fueron prácticamente nulas. Cokely et al. (2012) únicamente encontraron una correlación significativa con los factores de personalidad de EX y de AE y Teovanović et al. (2015) obtuvieron una asociación significativa con la dimensión de AE. Igualmente, Juanchich et al. (2016) encontraron una única asociación significativa entre el TRC y el factor EX.

Por otro lado, Cueva et al. (2015) llevaron a cabo un análisis de regresión para predecir el desempeño en el TRC mediante los cinco factores de personalidad. Los resultados mostraron que, únicamente, EX y Neuroticismo (N, *Neuroticism*) se asociaron significativa y negativamente con la probabilidad de responder correctamente a los ítems en el TRC. El análisis de regresión llevado a cabo por Alós-Ferrer et al. (2016) mostró que C y EX no fueron predictores independientes del desempeño en el TRC, aunque sí lo fueron N, A y AE. De modo que los más amigables, menos abiertos a la experiencia y menos neuróticos fueron los que mostraron una mayor probabilidad a responder más ítems correctamente en el test. Por último, Agnew y Harrison (2017) encontraron que las variables de personalidad que explicaron significativamente el desempeño en el TRC fueron EX y C.

Finalmente, otros estudios han examinado la relación entre el TRC y rasgos de personalidad relacionados con los BFF. Por ejemplo, se exploró la relación con la tendencia a posponer o aplazar las tareas, deberes u otras obligaciones por actividades más gratificantes (procrastinar) y se encontraron resultados muy variados al respecto. Frederick (2005) no encontró una correlación entre la autopercepción a procrastinar y las puntuaciones en el TRC. Sin embargo, Cesarini et al. (2012) encontraron una asociación entre ambas variables.

Por su parte, Aktaş et al. (2017) estudiaron la relación entre el TRC y la presencia de rasgos de psicopatía en los participantes. En este sentido, los resultados no fueron muy concluyentes. Los autores encontraron que el TRC se asoció con la psicopatía en uno de los tres estudios llevados a cabo.

3.1.2. Conductas Deshonestas

La relación entre el TRC y la comisión de conductas deshonestas (como p. ej. *cheating*) no ha sido totalmente esclarecida. Los estudios han arrojado resultados muy variados, desde una asociación nula entre ambas variables (Gino y Ariely, 2012; Lohse, Simon y Konrad, 2018) hasta que los sujetos más reflexivos (o los sujetos más intuitivos) pueden presentar una mayor probabilidad a comportarse deshonestamente (Fosgaard, Hansen y Piovesan, 2013; Ruffle y Tobol, 2017).

El estudio de Gino y Ariely (2012) mostró que el TRC no correlacionó significativamente con ninguna de las medidas de deshonestidad. Los resultados encontrados por Lohse et al. (2018) mostraron que el TRC no tuvo efectos sobre la deshonestidad en un juego de lotería. La puntuación obtenida en el TRC no correlacionó con la fracción de sujetos que falsificaron su respuesta frente al total de potenciales falsificadores ni tuvo efectos sobre la decisión final consciente de falsificar la respuesta. Sin embargo, el TRC sí se relacionó con la capacidad de reconocer la oportunidad de falsear la respuesta. También se halló que el tiempo de presión reduce la consciencia de la oportunidad de falsear las respuestas, por lo que se sugiere que la consciencia de la oportunidad de mentir podría ser considerada un determinante o una precondition de la decisión consciente de mentir.

Por otro lado, Fosgaard et al. (2013) examinaron si las diferencias entre hombres y mujeres en RC podrían explicar las diferencias de sexo en la comisión de conductas deshonestas. Los resultados mostraron que el TRC afectó positivamente a la probabilidad de engañar, de modo que los sujetos con puntuaciones más altas en el TRC mostraron una mayor probabilidad a engañar. Sin embargo, las diferencias de sexo en la comisión de conductas deshonestas no pudieron ser explicadas por las sistemáticas diferencias en RC ya que, una vez controlado por los efectos de las puntuaciones en el test, la variable sexo continuó siendo un predictor significativo de la comisión de conductas deshonestas.

Los resultados hallados por Ruffle y Tobol (2017) mostraron que, a medida que los sujetos obtienen mejores resultados en el TRC y en capacidad mental general, el resultado informado se aleja más del valor deseado o esperado. Los datos sugirieron que los sujetos con puntuaciones más bajas en el TRC y capacidad mental general podrían falsear sus respuestas al facilitar un valor más alto o próximo al valor deseado. Asimismo, el análisis de regresión confirmó que la RC y la capacidad cognitiva fueron predictores válidos del resultado reportado. Los pesos de las variables se mantuvieron constantes cuando se controló por el efecto de las variables de autoinforme de la honestidad. Los autores argumentan que los sujetos con mayor capacidad para inhibir la respuesta intuitiva en el test y dar paso a una respuesta más deliberada y correcta son los mismos que contemplan los costes y beneficios de mentir. Por lo mismo, les resultará más fácil resistirse a comportarse de forma deshonesto. Entre los costes de mentir estaría el hecho de enfrentarse a la realidad de ser descubierto de su engaño, con la consiguiente merma de su imagen social. Otra posible explicación sería que los sujetos más sofisticados cognitivamente son más rápidos aprendiendo y aplicando las normas sociales de honestidad, de forma que cuando se encuentran en una situación novedosa, los individuos necesitan interpretar la situación antes de decidir qué respuesta dar, siendo más capaces de detectar el objetivo real del experimento y optar por seguir las normas sociales.

3.2. Variables Demográficas

3.2.1. Sexo

En su estudio, Frederick (2005) encontró que los hombres obtuvieron puntuaciones promedio más altas que las mujeres en el TRC, lo que parece sugerir que las mujeres mantienen un menor control inhibitor en sus respuestas (Böckenholt, 2012). Posteriores estudios también han encontrado estas diferencias en función del sexo con independencia del tipo de muestra, país de realización del estudio y versión del test utilizado (Alós-Ferrer y Hügelschäfer, 2016; Brañas-Garza, et al., 2015; Brosnan et al., 2014; Campitelli y Gerrans, 2014; Guthrie et al., 2007; Nieuwenstein y van Rijn, 2012; Razmyar y Reeve, 2013; Willard y Norenzayan, 2017; Yilmaz y Saribay, 2016; Toplak et al., 2017).

Tres meta-análisis se han llevado a cabo con el objetivo de examinar esta cuestión. Brañas-Garza et al. (2015) integraron 118 estudios con una muestra acumulada de 44,558 sujetos. Los resultados mostraron que los hombres desempeñaron mejor cada ítem del TRC que las mujeres ($\beta = .113$ para el ítem del bate y la pelota; $\beta = .177$ para el ítem de las máquinas y $\beta = .197$ para el ítem de los nenúfares). Además, encontraron que es más probable que las mujeres no respondan ningún ítem correctamente a que lo hagan los hombres y, a su vez, es más probable que estos resuelvan los tres ítems correctamente a que lo hagan las mujeres. Igualmente encontraron que las diferencias en función del sexo persisten incluso cuando se ha controlado por los efectos de otras variables, como los incentivos económicos, el tipo de muestra (estudiantes vs.

no estudiantes), el formato del TRC (ordenador vs. lápiz y papel), la posición del TRC dentro del experimento, etc. Cueva y colaboradores (2015) integraron en total 8 estudios independientes ($N = 1,180$ sujetos) llevados a cabo por su grupo de investigación y reportaron el estadístico de contraste sobre las diferencias encontradas entre ambos grupos. Los resultados mostraron diferencias significativas entre hombre y mujeres en RC (1.12 vs. 0.58, $p < .001$). Finalmente, el meta-análisis de Primi et al. (2018; $K = 13$, $N = 2,536$) mostró un tamaño del efecto promedio observado de $d = .53$, IC 95% [.34, .72]. De acuerdo con los criterios de Cohen (1907), el tamaño de las diferencias fue medio.

Asimismo, Frederick (2005) examinó el tipo de respuestas erróneas cometidas por ambos grupos y encontró que cuando las mujeres fallan al responder el TRC suelen ofrecer en mayor medida la respuesta errónea intuitiva que los hombres, mientras que estos tienden a cometer una variedad más amplia de errores. Indicó que el ratio de los errores intuitivos, frente a otros errores, es mayor en mujeres que en hombres, sugiriendo que los hombres tienden a reflexionar más sus respuestas y están menos inclinados a responder de forma intuitiva. Otras investigaciones han encontrado resultados similares. Por ejemplo, Brosnan et al. (2014) encontraron que las mujeres puntuaron sistemáticamente más que los hombres en el TRC intuitivo. Pennycook, Cheyne, Koehler et al. (2015) hallaron una relación entre el TRC intuitivo y sexo, siendo las mujeres las que, en promedio, reprodujeron en mayor medida la respuesta intuitiva. Aunque, en este caso, los resultados también mostraron que las mujeres fueron más tendentes a dar otras respuestas erróneas (diferentes de las intuitivas). Finalmente, el estudio de Nieuwenstein y van Rijn (2012) ha sugerido que las diferencias entre hombres y mujeres en RC pueden ser debidas al tiempo invertido en la deliberación de las respuestas. Los resultados reflejaron que los hombres podrían invertir más tiempo procesando las respuestas que las mujeres, lo que les ayudaría a resolver correctamente más ítems.

Hasta el momento apenas se han investigado las causas de las diferencias en función del sexo en las puntuaciones en el TRC, sin embargo, múltiples razones podrían explicar estos resultados, incluyendo diferencias en la educación o crianza de ambos grupos (Bosch-Domènech et al., 2014). Una de las principales hipótesis que se ha tratado de testar es la relacionada con el contenido matemático del TRC. Estudios previos han sugerido que las diferencias en función del sexo en el TRC podrían ser debidas a las diferencias individuales en la capacidad cognitiva o la capacidad matemática. Como se ha indicado anteriormente, el TRC es un test que se apoya en la capacidad cognitiva y, especialmente, en la capacidad numérica. Aunque diversos meta-análisis han demostrado que no existen diferencias en función del sexo en el desempeño matemático (Else-Quest, Hyde y Linn, 2010; Lindberg, Hyde, Petersen, y Linn, 2010), otros han mostrado que las mujeres suelen experimentar mayores niveles de ansiedad y sentirse menos seguras de su capacidad para desempeñar tareas matemáticas (Else-Quest et al., 2010; Hyde, Fennema, Ryan, Frost y Hopp, 1990). Diversos estudios han encontrado una relación negativa entre la ansiedad

matemática y el desempeño en el TRC (Cokely et al., 2012; Primi et al., 2018; Primi et al., 2017; Skagerlund et al. 2018), siendo los efectos de la ansiedad matemática no solo directos sino también indirectos a través del conocimiento matemático (Morsanyi et al., 2014). Por lo tanto, el contenido numérico del TRC, bien sea a través de la capacidad, del conocimiento o de la ansiedad matemática, podría explicar las diferencias entre hombres y mujeres en las puntuaciones del TRC (Primi et al. 2015).

En base a ello, diversos estudios han tratado de examinar esta cuestión. Primi et al. (2015) llevaron a cabo un análisis de regresión para determinar el desempeño en el TRC. Los resultados mostraron que cuando se controló por capacidad numérica, las diferencias entre hombres y mujeres se redujeron significativamente, aunque el efecto de la variable sexo continuó siendo significativo. No obstante, en la segunda muestra examinada, cuando se controló por la capacidad numérica subjetiva, las diferencias entre ambos grupos desaparecieron. Estos resultados sugieren que la capacidad matemática subjetiva podría ser responsable de las diferencias de sexo en el TRC, mientras que la capacidad matemática podría serlo solo parcialmente.

Primi et al. (2018) examinaron la invarianza del TRC en función del sexo para comprobar si los ítems son adecuados para medir el rasgo latente en hombres y mujeres. Para ello, se verificó la estructura del test empleando un análisis de funcionamiento diferencial de los ítems de acuerdo con la teoría de respuesta al ítem. Los resultados mostraron la equivalencia de los ítems del test para medir el constructo en ambos grupos. El mismo rasgo latente fue evaluado con el TRC en hombres y mujeres y, por lo tanto, las diferencias en función del sexo pueden ser interpretadas como diferencias reales entre los grupos dentro del nivel del rasgo latente. A continuación, se comprobaron los efectos de la ansiedad y el razonamiento matemáticos sobre la relación entre la RC y el sexo. Los resultados mostraron que el efecto total del sexo en el TRC fue significativo y, una vez introducidas las variables mediadoras del modelo (ansiedad matemática y capacidad numérica), el efecto directo de la variable sexo dejó de ser significativo.

En conjunto, los resultados expuestos sugieren la existencia de diferencias en función del sexo en el desempeño en el TRC, siendo los hombres los que obtienen mejores resultados en el test. Asimismo, los datos evidencian que, las diferencias en función del sexo pueden desaparecer cuando se controla por los efectos de la ansiedad y la capacidad matemática. Aunque el TRC no es simplemente otra medida más de capacidad numérica, el contenido numérico de los ítems podría actuar como un importante determinante de su correcto desempeño. Consecuentemente, se ha tratado de crear nuevos ítems objetivos de RC que se apoyen en menor medida en la capacidad y conocimientos matemáticos (Thomson y Oppenheimer, 2016; Szasz et al., 2017) y las evidencias científicas han sugerido que hombres y mujeres no puntúan significativamente diferente en esos ítems (Thomson y Oppenheimer, 2016; Yilmaz y Saribay, 2017).

No obstante, diversos estudios muestran evidencias de que las diferencias en función del sexo en el TRC podrían ser independientes de la capacidad matemática. Frederick (2005), a pesar

de haber encontrado diferencias significativas entre hombres y mujeres en el TRC, no pudo demostrar diferencias en función del sexo en las restantes medidas cognitivas (SAT total, SAT verbal, ACT y WPT), salvo en la parte matemática de SAT. Con el objetivo de explicar las diferencias en función del sexo en el desempeño en el TRC llevó a cabo un análisis de regresión controlando los resultados por los efectos de las puntuaciones de SAT matemático. Los resultados reflejaron que las diferencias en función del sexo se mantuvieron estadísticamente significativas. Agnew y Harrison (2017) hallaron que las puntuaciones en el TRC continuaron siendo significativamente superiores para los hombres que para las mujeres una vez controlado los análisis por los efectos de las puntuaciones en GPA, capacidad cognitiva matemática y autoinforme de impulsividad. Por su parte, Welsh et al. (2013) mostraron que, los hombres puntuaron más alto que las mujeres en el TRC y, sin embargo, no pudieron demostrar diferencias en función del sexo en la capacidad matemática. Sin embargo, hallaron diferencias en función del sexo en la capacidad racional medida con REI, por lo que los autores sugirieron que estas diferencias podrían explicar las diferencias entre hombres y mujeres en la RC.

Asimismo, otra de las causas que podrían explicar las diferencias entre hombres y mujeres en el desempeño en el TRC son las diferencias entre ambos grupos en el ratio digital 2D:4D. El ratio 2D:4D es el ratio dimórfico sexual entre el largo del dedo índice y el dedo anular y se ha propuesto como el biomarcador que correlaciona negativamente con la exposición prenatal a la hormona esteroidea de la testosterona. Un índice 2D:4D mayor tiende a estar asociado con una menor exposición a la testosterona en el útero materno durante el primer trimestre del embarazo. Normalmente, son las mujeres las que están menos expuestas a esta hormona y, por extensión, tienden a presentar una mayor proporción 2D:4D que los hombres (Friedl, Neyse y Schmidt, 2018).

Bosch-Domènech et al. (2014) encontraron que el ratio 2D:4D fue superior en mujeres que en hombres para las dos manos, siendo significativamente mayor en la mano izquierda que en la derecha. Asimismo, encontraron una relación significativa y negativa entre el TRC y el ratio 2D:4D en mujeres, pero no en hombres. De modo que las mujeres que obtienen mejores resultados en el TRC presentan un ratio más pequeño y, por ende, mayor exposición a la testosterona en el útero. El análisis de regresión mostró que la capacidad matemática y la paciencia en el retraso de las gratificaciones explicaron varianza única del desempeño en el TRC, una vez controlado por el ratio 2D:4D y la variable sexo. Por lo que, ambas variables fueron determinantes independientes del desempeño en el TRC.

Del mismo modo, Cueva et al. (2015) encontraron que el valor promedio del ratio 2D:4D fue más bajo para los hombres y las mujeres con puntuaciones más altas en el TRC. Pero, sin embargo, no encontraron diferencias significativas en el valor promedio del ratio entre hombres con altas y bajas puntuaciones en el TRC ni entre las mujeres con altas y bajas puntuaciones en el test.

En síntesis, existen numerosas evidencias que indican que los hombres tienden a alcanzar sistemáticamente mejores resultados que las mujeres en el desempeño en el TRC. Además, existen indicios de que estas diferencias se deben a diferencias reales en el nivel de rasgo latente que mide el test y no a una diferente validez de los ítems para medir el constructo en ambos grupos. Asimismo, diversos estudios sugieren que la capacidad numérica, la ansiedad matemática y las diferencias en el ratio 2D:4D podrían mediar o explicar parcialmente las diferencias en función del sexo en el test. Estos datos sugieren que las variables determinantes de la RC en hombre y mujeres podrían ser diferentes.

En este sentido, Campitelli y Gerrans (2014) llevaron a cabo dos análisis de regresión independientes para predecir el desempeño en el test en función del sexo. Los resultados mostraron que las puntuaciones de *numeracy*, las tareas de sesgo de creencia en el razonamiento silogístico y el AOT fueron predictores significativos del TRC tanto en hombres como en mujeres. No obstante, la contribución de AOT en hombres fue mayor que en mujeres. Este análisis inicial sugiere que el TRC tiene un componente matemático y de pensamiento racional tanto en hombres como en mujeres y que la contribución de las disposiciones de pensamiento es débil, aunque significativo, especialmente en el caso de las mujeres. Además, los autores diseñaron 4 modelos matemáticos de análisis con el objetivo de testar el modelo de mejor ajuste en ambos grupos. Los modelos se basan en las principales teorías sobre qué mide el TRC. Así, (1) el primer modelo a testar fue un modelo nulo que asume que los sujetos desempeñan de forma similar el test. (2) El segundo modelo asume que el TRC solo mide capacidad numérica. (3) El tercer modelo defiende que el TRC mide capacidad matemática y pensamiento racional. Por último, (4) el cuarto modelo sostiene que el TRC mide capacidad matemática, pensamiento racional y la disposición a pensar activamente con la mente abierta. Los resultados mostraron que, para el caso de los hombres, el modelo de mejor ajuste fue el modelo de disposiciones mientras que, para el caso de las mujeres, el modelo de mejor ajuste fue el modelo racional. Los resultados sugieren que la disposición hacia un pensamiento de mente abierta no juega un rol significativo en el desempeño en el TRC en mujeres, pero sí en el desempeño de los hombres. El TRC mide capacidad matemática y pensamiento racional en mujeres mientras que mide capacidad matemática, pensamiento racional y disposición a pensar con la mente abierta en hombres. Los resultados apoyan la hipótesis de que el TRC no es una medida más de capacidad numérica y que las variables que determinan las puntuaciones en el test podrían ser diferentes en función del sexo.

3.2.2. Edad

La investigación sobre la relación entre el TRC y la edad refleja resultados muy diversos, siendo poco ilustrativos en cuanto a la real asociación entre ambas variables. En tendencia general, los estudios han sugerido que el nivel de RC de los sujetos es independiente de su edad cronológica (Aktaş et al., 2017; Narayanan y Moritz, 2015; Royzman et al., 2014, 2015; Sinayev

y Peters, 2015; Skagerlund et al., 2018; Toplak et al., 2017; Welsh et al., 2013; Willard y Norenzayan, 2017). Sin embargo, investigaciones sobre la impulsividad sugieren que la RC podría decrecer con la edad ya que, la función de autocontrol y autorregulación dependen de las capacidades cognitivas y, estas merman a lo largo de la vida (Finucane y Gullion, 2010). De este modo, la automaticidad y la consistencia de la experiencia demandada por las tareas diarias se incrementan con la edad con el único objetivo de reducir el consumo de recursos cognitivos. Así, diversos estudios han mostrado una asociación negativa entre las puntuaciones en el TRC y la edad (Albaity et al., 2014; Browne et al., 2014; Campitelli y Labollita, 2010; Ghazal et al., 2014; Juanchich et al., 2016; Koscielniak et al., 2016; Royzman et al., 2015; Yilmaz y Saribay, 2016, 2017).

Por ejemplo, Finucane y Gullion (2010) encontraron que, en promedio, los jóvenes (25-45 años) acertaron 2.5 ítems en el TRC, los jóvenes-mayores (65-74 años) 2.1 y los mayores-mayores (75-97 años) 1.7. Los resultados reflejaron un efecto significativo de la edad sobre la RC, demostrando que la edad puede influir en la capacidad de inhibir la respuesta impulsiva. Asimismo, Albaity et al. (2014) encontraron que los sujetos con 20 o menos años puntuaron, en promedio, 0.98 en el TRC, los sujetos de entre 21 y 30 años 0.83, los sujetos de entre 31 y 40 años 0.67 y los mayores de 41 años 0.63. Los resultados reflejaron una asociación negativa con la edad. Koscielniak et al. (2016) hallaron diferencias significativas entre las mujeres jóvenes-adultas y las mujeres adultas-mayores en el desempeño en el test, siendo las mujeres jóvenes-adultas las que alcanzaron mejores resultados. Finalmente, Besedeš, Deck, Sarangi y Shor (2012) hallaron una asociación entre la edad y los resultados de la toma de decisión del día a día. Emplearon una muestra de sujetos senior con edades comprendidas entre 60 y más de 80 años y mostraron que los individuos de mayor edad desempeñaron peor la tarea de toma de decisión que los seniors más jóvenes. Los resultados sugieren que las personas con una edad más avanzada podrían emplear diferentes heurísticos para resolver los problemas, siendo estas estrategias, frecuentemente, poco óptimas.

4. EVIDENCIAS DE VALIDEZ DE CRITERIO

Muy pocos estudios han examinado las evidencias de validez de criterio de la medida de RC. Algunas investigaciones han explorado la validez del TRC para predecir el desempeño académico, pero, sin embargo, la validez de este constructo para predecir criterios ocupacionales de interés apenas ha sido estudiada.

Teniendo en cuenta que el principal objetivo de esta tesis es examinar el constructo de la RC en el ámbito de la Psicología Organizacional, este apartado trata de sintetizar los principales resultados empíricos encontrados sobre su validez de criterio.

4.1. Desempeño Académico y Reflexibilidad Cognitiva

El desempeño académico está habitualmente expresado en términos de las calificaciones académicas o GPA (*Grade Point Average*; York, Gibson y Rankin, 2015). Aunque otras medidas de éxito académico como el grado de finalización de los estudios, el número de exámenes repetidos, la productividad investigadora o conductas académicas contraproductivas han sido utilizadas con propósitos de investigación, el GPA continúa siendo la medida de desempeño académico más ampliamente utilizada dentro de la literatura científica (Hughes y Douzenis, 1986; Kanoy, Wester, y Latta, 1989; Richardson, Abraham y Bond, 2012). En parte, esto puede ser debido a que el GPA es uno de los mejores indicadores del desempeño de los estudiantes ya que, se construye a través de múltiples evaluaciones procedentes de diversos profesores y materias y, además, engloba muestras de desempeño procedentes de un amplio rango de tiempo (Kuncel, 2003). Asimismo, es una de las variables más empleadas en procesos de selección académicos (acceso a estudios universitarios o de postgrado) y de contratación de empleados recién titulados (Kuncel, Hezlett y Ones, 2001; Strenze, 2007).

Los estudios que han examinado la validez de la RC para predecir el desempeño académico parecen indicar la existencia de una relación positiva entre ambos constructos (Borghans y Golsteyn, 2014). Por ejemplo, Insler y colaboradores (2015) encontraron una correlación de .31 ($N = 364$) entre RC y GPA. Toplak et al. (2014) hallaron una relación de .23 ($N = 160$) con GPA auto informado, que ascendió a .25 cuando la medida de RC empleada estuvo compuesta por 7 ítems (3 ítems originales y 4 adicionales). Thomson y Oppenheimer (2016) hallaron una asociación de .33 con el TRC original, mientras que, el TRC verbal se asoció en .27 ($N = 134$). Asimismo, Agnew y Harrison (2017) también encontraron una relación entre el TRC y las calificaciones académicas o GPA (coeficiente OLS = .068, $p = .05$). Sin embargo, Corgnet, Hernán-Gonzalez y Mateo (2015) encontraron una correlación prácticamente nula entre ambos constructos.

Del mismo modo, algunos estudios han comprobado la validez del TRC para predecir otros resultados académicos de interés. Así, se ha evidenciado que la RC es un válido predictor del logro matemático (Gómez-Chacón et al., 2014; Lindeman y Svedholm-Häkkinen, 2016), de las calificaciones en física (Lindeman y Svedholm-Häkkinen, 2016), de las calificaciones en la sección de matemáticas y lenguaje del examen de nivel A de alemán (Lohse, 2016) y de las calificaciones de exámenes de estadística (Primi et al., 2017). Sin embargo, no se pudo demostrar su validez para predecir la asistencia a clase (Primi et al., 2017).

Por otro lado, se ha examinado la validez añadida del TRC para predecir el éxito académico. Toplak et al. (2014) llevaron a cabo un análisis de regresión con el objetivo de predecir el GPA. Los resultados reflejaron que, únicamente la CMG y la composición de medidas de disposición de pensamiento explicaron las calificaciones. Además, tanto la CMG como las disposiciones de pensamiento explicaron varianza única del GPA (5.4% y 10.3%,

respectivamente). No obstante, el TRC desarrollado de 7 ítems no explicó varianza única del criterio. Por su parte, Gómez-Chacón et al. (2011) llevaron a cabo un estudio para examinar la relación entre el TRC y el logro matemático, entendido como la calificación final del curso de matemáticas. Además, examinaron el sistema de creencias de los estudiantes de secundaria sobre las matemáticas y el aprendizaje de los conocimientos matemáticos. Los resultados mostraron que el TRC correlacionó significativamente con la medida de creencias matemáticas y con el logro académico en esa disciplina; pero, el análisis de regresión reflejó que, una vez controlado por los efectos de la medida de creencias, el TRC no predijo significativamente el logro académico en matemáticas.

4.2. Desempeño Ocupacional y Reflexibilidad Cognitiva

Hasta el momento, la validez de la RC para predecir criterios organizacionales de interés no ha sido apenas investigada. Existe algún estudio en el que se ha explorado la relación entre este constructo y variables que se podrían considerar próximas a criterios organizacionales, pero que, en cualquier caso, no constituyen una medida de desempeño ocupacional (Corgnet, Hernán-González y Mateo, 2015; Narayanan y Moritz, 2015). A nuestro juicio, el único estudio que ha explorado previamente esta cuestión es el de Salgado et al. (2019). Estos autores encontraron una asociación positiva y significativa entre la RC y el desempeño ocupacional ($p = .32$). Para ello, emplearon, como instrumento de RC, una medida desarrollada por Salgado (2014a) y como medida *proxy* del desempeño ocupacional, la puntuación total obtenida en un *assessment center* (muestra de trabajo). Asimismo, los autores encontraron que, en conjunto, la RC y la CMG explicaron un 21.62% de la varianza del desempeño ocupacional y que ambas variables contribuyeron de forma independiente en la predicción del éxito ocupacional. Los resultados no solo sugieren la validez de la RC para predecir criterios organizacionales sino también evidencian su capacidad para añadir validez sobre otras variables individuales en la predicción de criterios ocupacionales.

Sin embargo, nuevos estudios son necesarios para demostrar la consistencia de los resultados obtenidos por estos autores. De este modo, uno de los objetivos de esta tesis ha sido explorar la capacidad de la RC para predecir el desempeño ocupacional y académico. Partiendo de que hay indicios de que la RC determina el desempeño académico y que existen meta-análisis que evidencian que el desempeño académico es un predictor del futuro desempeño ocupacional (Dye y Reek, 1988; Roth, BeVier, Switzer y Schippmann, 1996; Samson, Graue, Weinstein y Walberg, 1984), parece lógico suponer que nuevos estudios replicarán los resultados encontrados por Salgado et al. (2019). En cualquier caso, el estudio 3 mostrará nuevas evidencias al respecto.



PARTE **E**MPÍRICA





ESTUDIO 1.

Propiedades Psicométricas de una Versión Nueva del TRC

1. INTRODUCCIÓN

El principal objetivo de esta tesis es validar el constructo de la flexibilidad cognitiva (RC) en el ámbito de la Psicología del Trabajo. Sin embargo, para poder cumplir con ese propósito, se requiere de un instrumento que permita evaluar la RC de forma válida y fiable. Como previamente se ha señalado en el capítulo 2, el TRC original (Frederick, 2005; Kahneman y Frederick, 2002) cuenta con una serie de limitaciones que afectan a sus propiedades psicométricas y parten, en gran medida, del reducido número de ítems que componen el test. Entre estas limitaciones cabe señalar: 1) el test muestra una deficiente capacidad de discriminación entre sujetos con diferente nivel de rasgo latente (“efecto suelo”; Thomson y Oppenheimer, 2016; Toplak et al., 2014; Weller et al., 2013); 2) el test presenta una distribución de las puntuaciones sesgada con valores de curtosis y asimetría positivos y una forma binomial y no normal (Juanchich et al., 2016; Weller et al., 2013); 3) los coeficientes de fiabilidad reflejan valores no aceptables desde el punto de vista psicométrico (Aktaş et al., 2017; Drummond y Fischhoff, 2017; Teovanović et al., 2015) y 4) los ítems son altamente populares por la previa exposición a través de diversos métodos de divulgación (artículos, libros, seminarios, etc.; Alós-Ferrer et al., 2016; Thomson y Oppenheimer, 2016).

Consecuentemente, estudios recientes han propuesto nuevas versiones del TRC. Estas versiones han conseguido superar la última limitación expuesta; sin embargo, apenas han conseguido solucionar el resto. Por lo tanto, el principal objetivo de este estudio es examinar las propiedades psicométricas de una nueva versión del TRC (Salgado, 2014a).

Además, estudios previos también han indicado la existencia de diferencias en el desempeño en el TRC en hombres y mujeres, sugiriendo que ambos grupos se comportan de modo diferente en la realización del test (Campitelli y Gerrans, 2014; Frederick, 2005; Primi et al., 2018). Los resultados apuntan a que los hombres tienden a obtener mejores resultados que las mujeres (Brañas-Garza et al., 2015; Primi et al., 2018). De este modo, el segundo objetivo de este

estudio ha sido examinar las propiedades psicométricas del test en ambos grupos y comprobar posibles diferencias en su desempeño.

Finalmente, algunos investigadores han encontrado que determinados factores relacionados con la administración o aplicación del test pueden determinar las puntuaciones de RC. Entre estos factores se encuentra el formato de administración del test (papel y lápiz vs. ordenador). En este sentido, Brañas-Garza et al. (2015) encontraron que esta variable moderó las puntuaciones de la RC de los participantes, de modo que aquellos sujetos que respondieron al test en formato ordenador alcanzaron sistemáticamente mejores resultados que los que respondieron en papel y lápiz. Por lo tanto, este estudio también examinará si el formato de presentación del test afecta al desempeño en el mismo.

En base a lo expuesto con anterioridad, las hipótesis planteadas fueron:

Hipótesis 1a: *el TRC original (Frederick, 2005) refleja una distribución no normal y sesgada de las puntuaciones, con curtosis y asimetría positivas.*

Hipótesis 1b: *el nuevo TRC (Salgado, 2014a) muestra una distribución normal o aproximadamente normal de las puntuaciones, con valores de curtosis y asimetría más próximos a cero.*

Hipótesis 2: *los hombres obtienen mejores resultados que las mujeres en el TRC original y en el nuevo test.*

Hipótesis 3: *los sujetos que respondan a los test de RC (TRC original o TRC nuevo) en la versión del ordenador obtendrán mejores resultados que los que respondan en la versión de lápiz y papel.*

Hipótesis 4: *el TRC nuevo presenta un coeficiente de fiabilidad más elevado que el TRC original.*

2. MÉTODO

Muestra

La muestra empleada en este estudio se compuso de 1,367 sujetos, de los cuales el 64.16% fueron mujeres ($n = 877$). La edad promedio de los sujetos fue 21.72 años ($SD = 4.14$, rango = 17 a 58 años). Todos los participantes fueron estudiantes de la Universidad de Santiago de Compostela (USC) y el 15.56% de ellos eran postgraduados. El restante 84.44% estaban matriculados en alguna de las titulaciones de grado ofertadas por la USC. La Tabla 6 recoge las frecuencias de los alumnos según el área de estudio. Como se puede observar, el 57.64% de los participantes eran alumnos matriculados en alguna titulación de grado del área de las ciencias sociales, el 12.44% estaban matriculados en el momento de realización del estudio en alguna

titulación del área de ciencias experimentales o tecnológicas, el 22.75% pertenecían al área de las ciencias de la salud y el 4.32% al área de humanidades.

Procedimiento

Para llevar a cabo este estudio se reclutaron, a través de diferentes métodos, varias muestras de participantes entre los años 2015 a 2018. El 31% de los sujetos ($n = 423$) fueron reclutados en las aulas, mientras que el restante 69% ($n = 944$) fueron participantes voluntarios de estudios experimentales llevados a cabo por el Grupo de Investigación en Recursos Humanos y Rendimiento de la USC. En este último caso, los sujetos además de responder el TRC completaron otras pruebas con un propósito diferente al de este estudio.

Tabla 6.

Características de la Muestra (N = 1,367)

	<i>n</i>	%
Sexo		
Mujeres	877	64.16
Hombres	490	35.84
Edad		
≤ 20	660	48.28
21-23	457	33.43
> 23	247	18.07
No indicado	3	.22
Área de Estudio		
Ciencias sociales	788	57.64
Ciencias experimentales	170	12.44
Ciencias de la salud	311	22.75
Humanidades	59	4.32
No indicado	39	4.85
Lugar de realización del estudio		
Aula	423	31.00
Laboratorio	944	69.00
Formato de administración del test		
Papel y lápiz	733	53.62
Ordenador	634	46.38

Al 53.62% ($n = 733$) de la muestra se le administró el TRC en formato papel y lápiz, mientras que el 46.38% ($n = 634$) restante lo cubrió a través del ordenador. En los casos en los que el TRC fue administrado mediante el ordenador, cada uno de los ítems del test fue presentado de forma individual en la pantalla. Una vez completado el ítem, los participantes podrían pasar al

siguiente ítem presionando el botón “siguiente”. El programa requería haber respondido al ítem previo para completar el siguiente. Una vez presionado el botón “siguiente”, el programa informático no permitía a los participantes retroceder en las preguntas. Con independencia del modo de administración del test, se instruyó a los participantes en el carácter confidencial y anónimo del estudio y en el único propósito e interés en examinar los resultados de modo agregado.

Con el objetivo de calcular el coeficiente de fiabilidad test-retest se animó a los participantes de los estudios experimentales ($n = 944$) a participar de nuevo en un momento temporal posterior. Entre la realización del TRC en el momento temporal 1 y 2 transcurrieron, en promedio, 96.67 días ($SD = 170.90$). En este caso, la muestra se redujo a 295 sujetos, de los cuales 210 fueron mujeres. La edad promedio de la muestra fue de 21.89 ($SD = 4.12$) y el rango de edad abarcó desde los 17 a los 58 años.

Instrumento de Medida

Para medir la RC se administró el Test de Reflexibilidad Cognitiva desarrollado por Salgado (2014a). Este test está compuesto por 14 ítems entre los que se incluyen los tres originalmente diseñados por Frederick (2005; Kahneman y Frederick, 2002). Cada ítem consiste en un pequeño problema matemático que provoca una respuesta intuitiva inmediata. Los problemas se resuelven correctamente haciendo simples operaciones de aritmética. Los participantes tuvieron que contestar a cada ítem seleccionando una de las dos opciones de respuesta propuestas: (1) la respuesta inmediata e intuitiva pero incorrecta o (2) la respuesta deliberada y correcta. La Tabla 7 recoge los ítems del test y las opciones de respuesta.

Los tres primeros ítems son los tres desarrollados por Frederick y traducidos al español. El 6 es el problema análogo al ítem del bate y la pelota, es decir, el problema del plátano y el pastel. Los restantes son nuevos problemas desarrollados por Salgado (2014a). Algunos de ellos son diferentes versiones de los ítems originales (véase, por ejemplo, el ítem 2, 4 y 9).

Durante la realización del TRC, los participantes no dispusieron de límite temporal para resolver los problemas. Por cada acierto los participantes recibieron un punto y por cada error 0 puntos, por lo tanto, las puntuaciones finales de cada participante pueden oscilar entre 0 y 14. Puntuaciones altas en esta medida indican mayor nivel de RC.

Tabla 7.
Ítems del TRC Desarrollado por Salgado (2014a)

N.º ítem	Denominación abreviada	Ítem	Opciones respuesta
1	Bate y pelota	Un bate y una pelota cuestan 1.10 euros. El bate cuesta un euro más que la pelota. ¿Cuánto cuesta la pelota?	R. intuitiva: 10 cént. R. correcta: 5 cént.
2	Máquinas	Si 5 máquinas tardan 5 minutos en hacer 5 aparatos, ¿cuánto tiempo tardan 100 máquinas en hacer 100 aparatos?	R. intuitiva: 100 min R. correcta: 5 min.
3	Nenúfares	En un lago, hay una zona de hojas de nenúfar. Cada día, la zona se duplica en tamaño. Si se tardan 48 días para que la zona cubra todo el lago, ¿cuánto tiempo haría falta para que la zona de nenúfares cubra la mitad del lago?	R. intuitiva: 24 días R. correcta: 47 días
4	Veleros	18 veleros navegando a 18 nudos tardan 18 días en ir del puerto A al puerto B. ¿Cuánto tardarán 36 veleros a 36 nudos?	R. intuitiva: 36 días R. correcta: 9 días
5	Veinteava parte	La veinteava (1/20) parte de veinte piezas de una cosa es igual a 2. ¿A cuánto será igual la cuarentava (1/40) parte de cuarenta piezas de la misma cosa?	R. intuitiva: 4 R. correcta: 2
6	Plátano y pastel	Un plátano y un pastel cuestan 37 céntimos. El plátano cuesta 13 céntimos más que el pastel. ¿Cuánto cuesta el pastel?	R. intuitiva: 24 cént. R. correcta: 12 cént.
7	Taxista	La probabilidad de que un taxista en una gran capital tenga un accidente al mes es del 10%. ¿Cuál será la probabilidad de que tenga 1 accidente en dos meses seguidos, por tanto, dos accidentes en total?	R. intuitiva: 20% R. correcta: 1%
8	Posibilidad de aprobar	En un colegio, los estudiantes asignados a la clase A aprueban todos, los asignados a la clase B aprueban la mitad y los asignados a la clase C aprueban uno de cada tres. ¿Qué posibilidades de aprobar tiene un nuevo estudiante antes de ser asignado a la clase A, B o C?	R. intuitiva: No se sabe antes de asignarlo R. correcta: 60%

Continúa

Tabla 7.
Continuación

N.º ítem	Denominación abreviada	Ítem	Opciones respuesta
9	Aviones	1,000 aviones tardan 1,000 minutos en transportar 1,000 pasajeros de la ciudad A a la ciudad B que está a más de 10,000 kilómetros de distancia. ¿Cuántos minutos tardarán 10,000 aviones en transportar 10,000 pasajeros?	R. intuitiva: 10,000 min. R. correcta: 1,000 min.
10	Coste pieza	Si el coste de 1 pieza por una máquina es de 100 céntimos y el coste de 50 piezas, por 50 máquinas es de 50 céntimos, ¿cuál será el coste de 100 piezas realizadas por 100 máquinas?	R. intuitiva: 25 cént. R. correcta: 1 cént.
11	Máquinas coste pieza	Si una máquina produce una pieza por 100 céntimos y 50 máquinas producen 50 piezas por 50 céntimos, ¿cuántas máquinas se necesitarán para que el precio de las piezas sea 12.5 céntimos?	R. intuitiva: 125 R. correcta: 200
12	Hombres altos	En una ciudad, de cada 4 personas 3 son hombres y 2 de cada 4 hombres son altos. ¿Qué posibilidades hay de que un día al salir a la calle, la primera persona que veamos sea un hombre alto?	R. intuitiva: 75% R. correcta: 37.5%
13	Sillas	¿De cuántos modos diferentes pueden sentarse 6 personas en una mesa con 6 sillas?	R. intuitiva: 36 R. correcta: 720
14	X + Y	$X + Y = 110$. $X = 100 + Y$. ¿Cuánto vale Y?	R. intuitiva: 10 R. correcta: 5

Análisis Estadísticos

Se llevaron a cabo 4 análisis estadísticos diferentes:

1. Análisis de los estadísticos descriptivos. Se procedió a examinar los estadísticos descriptivos del TRC original de Frederick (2005; en adelante TRC-3), los nuevos 11 ítems incorporados por Salgado (2014a; en adelante TRC-11) y el TRC completo de 14 ítems (3 del TRC original y los 11 incorporados por Salgado; en adelante TRC-14). Entre los estadísticos descriptivos se examinó la media, la desviación típica, la mediana, la moda, los cuartiles, la curtosis y la asimetría de las puntuaciones.
2. Frecuencias. Asimismo, se examinaron las frecuencias de ítems respondidos correctamente y la distribución de las puntuaciones. Estos estadísticos se calcularon para el total de la muestra y para la muestra de hombres y de mujeres por separado.
3. Análisis de varianza. Se examinaron las diferencias en función del sexo en el desempeño en cada uno de los ítems y en cada TRC concreto (TRC-3, TRC-11 y TRC-14). También se examinaron las diferencias en función del formato de administración del test (ordenador vs. lápiz y papel).
4. Análisis factorial. Se examinó la estructura interna del TRC de 14 ítems llevando a cabo un análisis factorial con el programa FACTOR (Ferrando y Lorenzo-Seva, 2017).

Lloret-Segura, Ferreres-Traver, Hernández-Bacza y Tomás-Marco (2014) y Ferrando y Lorenzo-Seva (2014) han indicado tres aspectos a tener en cuenta para determinar la matriz de correlaciones que se debe emplear en el análisis factorial. (1) El primero de ellos está asociado con la naturaleza de las variables o el número de categorías de respuesta de los ítems. El modo de corrección del test implica la dicotomización de las variables observadas (es decir, los ítems), de modo que las respuestas son evaluadas mediante datos ordenados en dos categorías (1 y 0). En los casos en que dos variables son categóricas (bien sean dicotómicas o politómicas), lo correcto es que los coeficientes de correlación aplicados sean las correlaciones tetracóricas, cuando se trate de variables dicotómicas, o las correlaciones policóricas, cuando estemos ante variables discretas con más de dos categorías de respuesta como, por ejemplo, escalas de tipo Likert (Ferrando y Lorenzo-Seva, 2014; Lloret-Segura et al., 2014). (2) El segundo aspecto que se debe valorar es el tamaño de la muestra. Se recomienda que el tamaño de la muestra sea igual o superior a 200 sujetos, ya que, en caso contrario, las correlaciones policóricas o tetracóricas pueden no resultar estables y ser menos precisas que las correlaciones de Pearson (Lloret-Segura et al., 2014). (3) El tercer aspecto a tener en cuenta está asociado con la distribución de las puntuaciones. La recomendación de los autores sobre este punto indica que, si nos encontramos ante muestras sesgadas o no distribuidas normalmente, lo

correcto sería emplear coeficientes de correlación tetracóricos o policóricos. Se considerará una distribución normal de las puntuaciones cuando los niveles de curtosis y asimetría sean moderados; y esto dependerá, a su vez, del criterio empleado. Un criterio exigente emplea valores en el rango -1 y $+1$, de modo que valores de asimetría y curtosis entre -1 y $+1$ indicarían una distribución normal o aproximadamente normal de las puntuaciones (Lloret-Segura et al., 2014). Por tanto, como recomendación general, las correlaciones de Pearson quedan limitadas a los casos en que los ítems sean continuos o, si no lo son, tengan 5 o más alternativas de respuesta y distribuciones aproximadamente normales. En los restantes casos, los autores aconsejan usar las correlaciones tetracóricas o policóricas según cada caso, salvo que la muestra de sujetos sea pequeña (Lloret-Segura et al., 2014; Ferrando y Lorenzo-Seva, 2014).

En base a esto, para el análisis factorial de nuestro estudio hemos empleado el programa FACTOR (Ferrando y Lorenzo-Seva, 2017) ya que permite calcular la matriz de correlaciones entre los ítems mediante coeficientes de correlación tetracóricos que son los pertinentes para el caso que nos ocupa. Concretamente, hemos empleado un análisis factorial exploratorio mediante el procedimiento de implementación óptima de un análisis paralelo para determinar el número de factores latentes (Timmerman y Lorenzo-Seva, 2011). Asimismo, se han estimado los indicadores de adecuación de la matriz de correlaciones de Kaiser, Meyer y Olkin (KMO) y el estadístico de Bartlett. Por último, se han calculado el índice de bondad de ajuste comparativo (*CFI*) y la raíz del residuo cuadrático promedio de aproximación (*RMSEA*) sobre el ajuste estadístico del modelo.

5. Cálculo de los coeficientes de fiabilidad. Se estimaron cuatro tipos diferentes de coeficientes de fiabilidad por cada TRC para la muestra total y para hombres y mujeres por separado. Concretamente, se calcularon dos tipos de coeficientes de fiabilidad por consistencia interna: (1) el Kuder-Richardson 20 (*KR20*), variante del alfa de Cronbach, y (2) el coeficiente EAP, calculado con el programa FACTOR, que parte del estimador bayesiano de la distribución posterior (Bock y Mislevy, 1982; Ferrando y Lorenzo-Seva, 2016). Asimismo, se calculó el coeficiente de estabilidad temporal o test-retest que estima el error cometido como consecuencia del paso del tiempo y el coeficiente de fiabilidad de equivalencia y estabilidad (*CES*) que estima la magnitud del error de medida producida por 3 fuentes diferentes de error: errores por respuesta aleatoria, errores específicos del factor y errores por el paso del tiempo (Schmidt, Le y Ilies, 2003). Para el cálculo de los coeficientes de fiabilidad test-retest del TRC original se emplearon las correlaciones policóricas calculadas con el programa FACTOR, ya que se trata de una variable discreta con 4 categorías de puntuaciones (0, 1, 2 o 3) y su distribución tiende a ser sesgada. Para

calcular el coeficiente de fiabilidad *CES* se emplearon las fórmulas 14 y 17a propuestas por Schmidt et al. (2003) para el caso de que el número de ítems del test sea par o impar, respectivamente. Este coeficiente no se pudo calcular para el TRC original por el limitado número de ítems que lo componen.

3. RESULTADOS

Estadísticos Descriptivos del TRC

La Tabla 8 muestra los estadísticos descriptivos del TRC-3, TRC-11 y TRC-14 para el total de la muestra y para las muestras de hombres y mujeres por separado.

Como se puede observar, el TRC-3 está compuesto por 3 ítems donde el rango de posibles puntuaciones oscila entre 0 y 3. El TRC-11 se compone de 11 ítems donde el rango de posibles puntuaciones varía entre 0 y 11 puntos y, finalmente, el TRC-14 se compone de 14 ítems y el rango de posibles puntuaciones puede fluctuar entre 0 y 14. La puntuación media de la muestra total fue de 1.42 ($SD = 1.08$), 6.41 ($SD = 2.59$) y 7.83 ($SD = 3.38$) para el TRC-3, el TRC-11 y el TRC-14, respectivamente. El valor de la mediana fue igual a 1 para el TRC-3, 7 para el TRC-11 y 8 para el TRC-14.

Por su parte, la moda o el valor al que le corresponde la mayor frecuencia ha sido 1 para el TRC-3, 8 para el TRC-11 y 11 para el TRC-14. De modo que, de las posibles puntuaciones que los participantes han podido obtener en el TRC-3, un porcentaje superior de la muestra ha alcanzado 1 ítem correctamente, en el TRC-11 8 ítems y en el TRC-14 11 ítems. Las Tablas 9 a 11 muestran las frecuencias para cada rango de puntuaciones y las Figuras 1 a 3 representan la distribución general de las puntuaciones de cada uno de los instrumentos de RC.

Asimismo, se puede observar que las tres distribuciones son asimétricas (.13, -.26 y -.16, para el TRC-3, TRC-11 y TRC-14, respectivamente), es decir, la proporción de sujetos con puntuaciones superiores o inferiores a la media no es homogénea. No obstante, el valor de asimetría del TRC-3 es positivo mientras que el del TRC-11 y el del TRC-14 son negativos. Esto indica que, para el TRC-3, la proporción de sujetos con puntuaciones inferiores a la media es mayor que la proporción de sujetos con puntuaciones superiores a la media. Sin embargo, para el TRC-11 y para el TRC-14, se da el patrón opuesto: la proporción de sujetos con puntuaciones superiores a la media es mayor que la proporción de sujetos con puntuaciones inferiores a la media.

Tabla 8.
Estadísticos Descriptivos del TRC

	TRC-3			TRC-11			TRC-14		
	N Total	Hombres	Mujeres	N Total	Hombres	Mujeres	N Total	Hombres	Mujeres
Tamaño de la muestra	1,367	490	877	1,367	490	877	1,367	490	877
N.º ítems	3	3	3	11	11	11	14	14	14
Rango posibles puntuaciones	0-3	0-3	0-3	0-11	0-11	0-11	0-14	0-14	0-14
Rango puntuaciones alcanzadas	0-3	0-3	0-3	0-11	0-11	0-11	0-14	0-14	0-14
Puntuación Promedio	1.42	1.72	1.24	6.41	7.37	5.88	7.83	9.10	7.12
Media									
Mediana	1	2	1	7	8	6	8	10	7
Moda	1	3	1	8	9	5	11	11	6
Medidas de variación									
Desviación Estándar	1.08	1.08	1.05	2.59	2.45	2.51	3.38	3.22	3.26
Asimetría	.13	-.25	.33	-.26	-.71	-.06	-.16	-.60	.03
Curtosis	-1.26	-1.22	-1.09	-0.72	-0.04	-0.71	-0.85	-0.32	-0.80
Cuartiles									
25	0	1	0	5	6	4	5	7	5
50	1	2	1	7	8	6	8	10	7
75	2	3	2	8	9	8	11	12	10

Nota. TRC-3 = TRC original (Frederick, 2005); TRC-11 = nuevos 11 ítems incorporados por Salgado (2014a); TRC-14 = ítems Frederick (2005) y 11 ítems Salgado (2014a); N total = muestra total.

Asimismo, los valores de curtosis encontrados han sido -1.26 para el TRC-3, -0.72 para el TRC-11 y -0.85 para el TRC-14. El valor negativo de la curtosis indica que la distribución de las puntuaciones es platicúrtica, es decir, el grado de apuntamiento de la distribución de frecuencias es aplastada, siendo la distribución del TRC-3 más aplastada que las distribuciones del TRC-11 y del TRC-14. Como se puede observar en la Figura 1, el grado de aplastamiento de la distribución del TRC-3 representa una forma rectangular, lo que indica una distribución binomial de las puntuaciones. No obstante, valores de asimetría y curtosis inferiores al valor absoluto de 1, apuntan hacia una distribución aproximadamente normal (Lloret-Segura et al., 2014), lo que indica que, tanto el TRC-11 como el TRC-14, aunque muestran distribuciones sesgadas, estas se aproximan a la normalidad (véanse las Figuras 1, 2 y 3). En conjunto, estos datos apoyan las Hipótesis 1a y 1b.

Por otro lado, se han hallado los cuartiles de la distribución de las puntuaciones. Para el TRC-3, el 25% de la muestra obtiene puntuaciones iguales a 0, el 50% iguales o inferiores a 1 y el 75% puntuaciones iguales o inferiores a 2. Para el TRC-11, el 25% de la muestra obtiene puntuaciones iguales o inferiores a 5, el 50% iguales o inferiores a 7 y el 75% iguales o inferiores a 8. Por último, para el TRC-14, el 25% de la muestra obtiene puntuaciones iguales o inferiores a 5 puntos, el 50% iguales o inferiores a 8 y el 75% puntuaciones iguales o inferiores a 11.

Tabla 9.
*Frecuencia de las Puntuaciones del TRC-3
(Muestra Total)*

Puntuación	Frecuencias
0	344
1	402
2	329
3	292

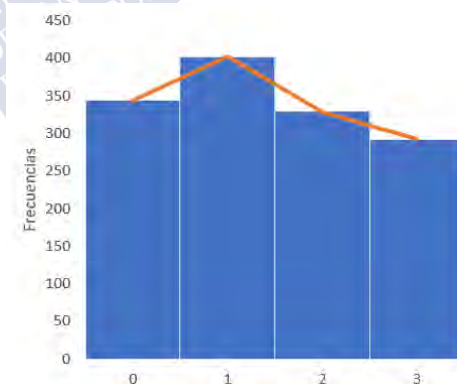


Figura 1.
*Distribución de las Puntuaciones del TRC-3
(Muestra Total)*

Tabla 10.
Frecuencia de las Puntuaciones del TRC-11
(Muestra Total)

Puntuación	Frecuencias
0	11
1	32
2	65
3	101
4	128
5	166
6	160
7	182
8	187
9	171
10	115
11	49

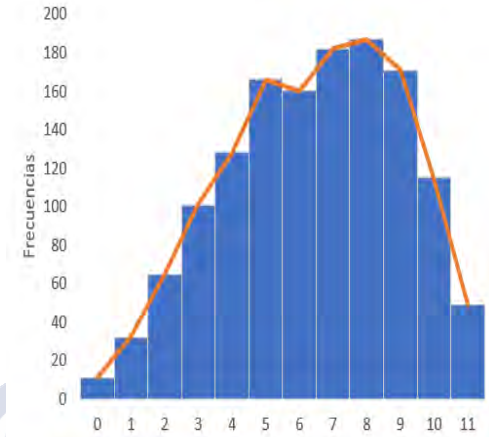


Figura 2. *Distribución de las Puntuaciones del TRC-11 (Muestra Total)*

Tabla 11.
Frecuencia de las Puntuaciones del TRC-14
(Muestra Total)

Puntuación	Frecuencias
0	11
1	27
2	55
3	69
4	89
5	115
6	141
7	119
8	134
9	124
10	113
11	157
12	100
13	80
14	33

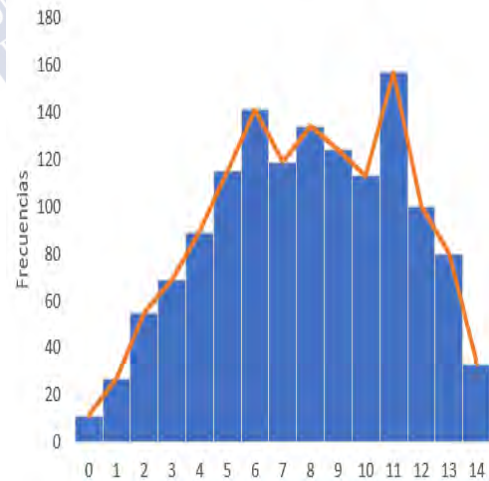


Figura 3. *Distribución de las Puntuaciones del TRC-14 (Muestra Total)*

Finalmente, la Tabla 12 recoge el porcentaje de personas que han respondido correctamente a cada uno de los ítems en el TRC. Como se puede observar, menos del 50% de la muestra ha respondido correctamente a los ítems 1, 3, 8, 11, 13 y 14; lo que indica que estos ítems, en promedio, son más difíciles de resolver que los restantes. El ítem 1 (el bate y la pelota) presenta un mayor grado de dificultad con un porcentaje de aciertos del 38%. Por orden de dificultad, le siguen el ítem 13 (con un 42.7% de respuestas correctas), el 3 y 14 (con un 43.5% de aciertos) y el 11 y 8 (con un 46.2 y 46.7% de aciertos, respectivamente).

No obstante, los ítems con una mayor proporción de aciertos y, por ende, los más fáciles, han sido el 9 (ítem de los aviones) con un 78.90% de aciertos, el 12 (ítem de los hombres altos) con un 74% de aciertos y el 4 (ítem de los veleros) con un 73.10% de aciertos. Los restantes ítems muestran porcentajes de acierto en torno al 50 y 63%. Estos resultados sugieren que, el TRC-14 podría ser un buen test para discriminar entre los diversos niveles de RC.

Tabla 12.

Porcentajes de Respuestas Correctas y Diferencias en Función del Sexo

	N Total	Hombres	Mujeres	<i>d (F)</i>
Ítem 1	38.0	47.1	33.0	.29 (27.53*)
Ítem 2	60.1	68.4	55.5	.25 (21.92*)
Ítem 3	43.5	56.9	35.9	.43 (58.89*)
Ítem 4	73.1	82.7	67.7	.34 (36.48*)
Ítem 5	57.4	67.6	51.8	.33 (32.75*)
Ítem 6	62.0	71.6	56.6	.31 (30.96*)
Ítem 7	53.5	68.2	45.3	.47 (69.51*)
Ítem 8	46.7	53.3	43.1	.20 (13.15*)
Ítem 9	78.9	84.7	75.6	.22 (15.76*)
Ítem 10	63.2	63.3	63.2	0
Ítem 11	46.2	55.9	40.8	.30 (29.4*)
Ítem 12	74.0	87.3	66.5	.49 (74.88*)
Ítem 13	42.7	51.8	37.6	.29 (26.40*)
Ítem 14	43.5	50.8	39.5	.24 (16.69*)
<i>N</i>	1,367	490	877	

Nota. * $p < .001$

Resultados en Función del Sexo

Respecto a los estadísticos descriptivos de la muestra de hombres y de mujeres por separado, los resultados mostraron que, en el TRC-3, la puntuación media fue de 1.72 ($SD = 1.08$) y 1.24 ($SD = 1.05$), respectivamente; en el TRC-11 fue de 7.37 ($SD = 2.45$) para la muestra de hombres y 5.88 ($SD = 2.51$) para la muestra de mujeres; y en el TRC-14 la puntuación promedio de los hombres fue 9.10 ($SD = 3.22$) y la de las mujeres 7.12 ($SD = 3.26$). Como se puede observar, en los tres TRC las puntuaciones promedio de las mujeres han sido más bajas que la de los hombres, cumpliéndose, así, la Hipótesis 2 formulada. Además, estas diferencias han sido estadísticamente significativas en los tres casos, con tamaños del efecto moderados de $d = .45$, $d = .60$ y $d = .61$ para el TRC-3, TRC-11 y TRC-14, respectivamente, de acuerdo con los criterios de Cohen (1977).

Consecuentemente, los valores de la mediana han sido mayores para los hombres que para las mujeres tanto en el TRC-3 ($Me_H = 2$ y $Me_M = 1$) como en el TRC-11 ($Me_H = 8$ y $Me_M = 6$) y TRC-14 ($Me_H = 10$ y $Me_M = 7$). No obstante, la moda ha sido 1 para la muestra de mujeres y 3 para la muestra de hombres en el TRC-3. Los datos indican que un porcentaje mayor de mujeres responden correctamente a únicamente un ítem del TRC-3, mientras que, en el caso de los hombres, lo hacen a los 3 ítems. En el TRC-11, la moda fue igual a 5 para las mujeres y 9 para los hombres, lo que indica que, de todas las posibles puntuaciones, un porcentaje mayor de mujeres ha respondido correctamente 5 ítems, mientras que en el caso de los hombres la frecuencia más alta se concentra en la puntuación 9. En el TRC-14, la moda fue igual a 9 para el grupo de mujeres y 11 para los hombres. Las Tablas 13 a 18 muestran las frecuencias para cada rango de puntuaciones en hombres y mujeres y las Figuras 4 a 9 representan gráficamente la distribución de las puntuaciones de los tres instrumentos de RC.

Nuevamente, se puede observar como las distribuciones de las puntuaciones del TRC-3, TRC-11 y TRC-14 para los hombres y las mujeres son asimétricas. La asimetría fue de $-.25$ para los hombres y de $.33$ para las mujeres en el TRC-3. Los datos indican una asimetría negativa para el caso de los hombres y positiva para el caso de las mujeres. Por lo que, la distribución de los hombres recoge muchas más frecuencias de valores altos en la variable que valores bajos; mientras que, la distribución de las mujeres muestra el patrón opuesto. Esto, a su vez, indica que las mujeres obtienen, en promedio, peores resultados que los hombres en el desempeño en el TRC-3. En el TRC-11, los valores de asimetría han sido $-.71$ y $.06$ para hombres y mujeres, respectivamente. En este caso, se puede observar como la distribución de las puntuaciones en la muestra de las mujeres refleja una distribución prácticamente simétrica, mientras que, en la muestra de hombres, la distribución sigue estando sesgada, con mayor frecuencia de valores altos de la variable.

Finalmente, para el TRC-14, los resultados muestran el mismo patrón que en el TRC-11. Los valores de asimetría muestran una distribución asimétrica negativa para la muestra de hombres (-.60) y una distribución simétrica en la de mujeres (.03). Asimismo, los valores de curtosis encontrados en el TRC-3, TRC-11 y TRC-14 reflejan valores negativos tanto en la muestra de hombres como de mujeres. Los datos manifiestan una distribución de las puntuaciones platocúrtica o con un grado de apuntamiento de la distribución de frecuencias aplastada. La distribución de las puntuaciones de la muestra de las mujeres es más aplastada que la de los hombres, salvo para el caso del TRC-3, donde el patrón se invierte, a pesar de haber diferencias muy pequeñas. En el caso del TRC-11, el valor de curtosis de la distribución de los hombres es muy próxima a 0 (-0.04), lo que indica que el grado de aplastamiento de la distribución es mesocúrtica o similar al grado de aplastamiento de la distribución normal. En cualquier caso, la distribución del TRC-3 muestra un valor mayor de curtosis que las otras dos distribuciones, tanto en la muestra de hombres como en la de mujeres. Los resultados de la curtosis y asimetría indican nuevamente que el TRC-11 y el TRC-14 muestran distribuciones aproximadamente normales, mientras que el TRC-3 refleja una distribución sesgada de las puntuaciones para las muestras de hombres y mujeres (véanse las Figuras 4 a 9).

Tabla 13.

Frecuencia de las Puntuaciones del TRC-3 (Hombres)

Puntuación	Frecuencias
0	83
1	121
2	134
3	152

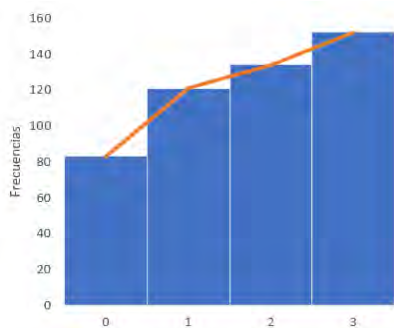


Figura 4. *Distribución de las Puntuaciones del TRC-3 (Hombres)*

Tabla 14.

Frecuencia de las Puntuaciones del TRC-3 (Mujeres)

Puntuación	Frecuencias
0	261
1	281
2	195
3	140

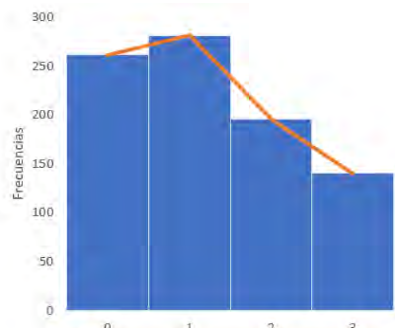


Figura 5. *Distribución de las Puntuaciones del TRC-3 (Mujeres)*

Tabla 15.
Frecuencia de las Puntuaciones del TRC-11 (Hombres)

Puntuación	Frecuencias
0	4
1	4
2	15
3	18
4	26
5	41
6	48
7	64
8	77
9	98
10	64
11	31

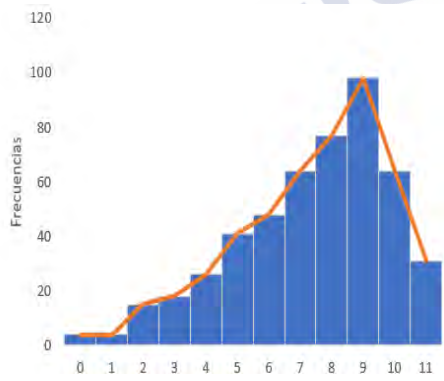


Figura 6. Distribución de las Puntuaciones del TRC-11 (Hombres)

Tabla 16.
Frecuencia de las Puntuaciones del TRC-11 (Mujeres)

Puntuación	Frecuencias
0	7
1	28
2	50
3	83
4	102
5	125
6	112
7	118
8	110
9	73
10	51
11	18

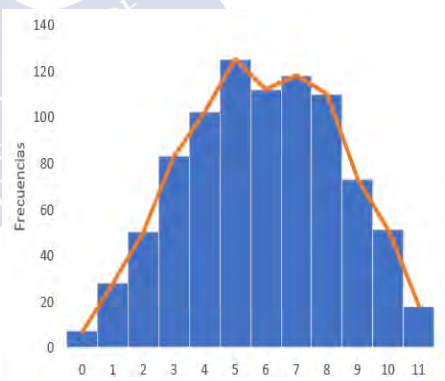


Figura 7. Distribución de las Puntuaciones del TRC-11 (Mujeres)

Tabla 17.
Frecuencia de las Puntuaciones del TRC-14 (Hombres)

Puntuación	Frecuencias
0	4
1	3
2	12
3	13
4	17
5	22
6	38
7	37
8	43
9	48
10	54
11	71
12	55
13	53
14	20

Tabla 18.
Frecuencia de las Puntuaciones del TRC-14 (Mujeres)

Puntuación	Frecuencias
0	7
1	24
2	43
3	56
4	72
5	93
6	103
7	82
8	91
9	76
10	59
11	86
12	45
13	27
14	13

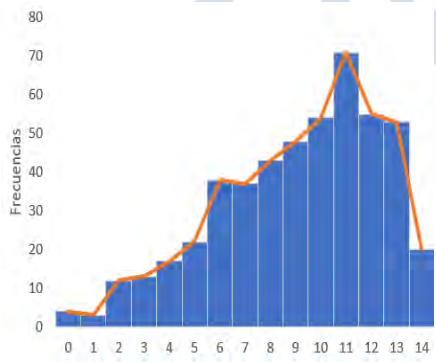


Figura 8. *Distribución de las Puntuaciones del TRC-14 (Hombres)*

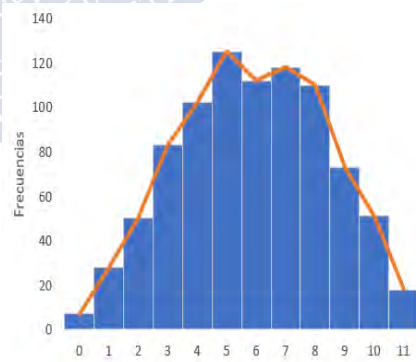


Figura 9. *Distribución de las Puntuaciones del TRC-14 (Mujeres)*

Por otro lado, se han hallado los percentiles de la distribución de puntuaciones y se ha obtenido que, en el TRC-3, el 75% de las mujeres obtuvieron puntuaciones iguales o inferiores a 2, mientras que los hombres obtienen puntuaciones iguales o inferiores a 3. En el TRC-11, el 75% de las mujeres obtuvieron puntuaciones iguales o inferiores a 8 y en el caso de los hombres iguales o inferiores a 9. Por último, en el TRC-14, el 75% de las mujeres alcanzaron puntuaciones iguales o inferiores a 10 puntos y en la muestra de hombres alcanzaron puntuaciones iguales o inferiores a 12 puntos. En su conjunto, estos datos nos indican que los hombres alcanzan, en promedio, mejores resultados en los tres instrumentos de RC que las mujeres.

Finalmente, se ha examinado el porcentaje de hombres y de mujeres que han respondido correctamente a cada uno de los ítems en el TRC y las diferencias en su desempeño en función del sexo (véase Tabla 12). En promedio, los resultados muestran que una proporción mayor de hombres que de mujeres responden correctamente cada ítem del TRC. Las diferencias encontradas han sido significativas y los tamaños del efecto oscilaron entre $d = .20$, en el caso del ítem 8 y $d = .49$, en el caso del 12; a excepción del 10 (ítem del coste por pieza) en el que no se han encontrado diferencias entre sexos ($d = 0$). Como se puede observar, el ítem con menor proporción de respuestas correctas ha sido el 1, donde un 47% de los hombres consiguieron responder correctamente el ítem frente a un 33% de mujeres. Para el resto de los ítems, al menos el 50% de los hombres han conseguido responder correctamente a todos ellos. En el caso de las mujeres, menos del 50% han conseguido responder correctamente los ítems 3, 7, 8, 11, 13 y 14.

Formato de Administración del TRC (Lápiz y Papel vs. Ordenador)

La muestra empleada en este estudio se puede dividir en dos submuestras en función del formato de administración del test (papel y lápiz vs. ordenador). Por lo que, se examinó si el formato de administración del test produce algún efecto sobre el desempeño de los participantes. La Tabla 19 recoge los resultados encontrados para el total de la muestra y las submuestras de hombres y mujeres en función del TRC empleado. Los datos reflejan que, con independencia de la muestra examinada, los participantes obtuvieron mejores resultados en los tres test cuando respondieron los ítems en el ordenador que cuando lo hicieron en papel y lápiz. No obstante, las diferencias solamente fueron significativas en el TRC-3 para la muestra total y para la muestra de hombres ($d = -.16$ y $d = -.21$), y en el TRC-14 para la muestra total ($d = -.12$). En todos los casos, los tamaños del efecto fueron pequeños, de acuerdo con los criterios de Cohen (1977). En el TRC-11, no se encontraron diferencias significativas en ninguna de las tres muestras. Estos resultados sugieren que el formato de administración del test podría no moderar el desempeño de los participantes en los nuevos ítems desarrollados por Salgado (2014a). De este modo, los resultados encontrados apoyan parcialmente la Hipótesis 3.

Tabla 19.

Medias, Desviaciones Típicas y Diferencias en Función del Formato de Administración del TRC

	TRC-3			TRC-11			TRC-14		
	N Total	H	M	N Total	H	M	N Total	H	M
Papel y lápiz									
\bar{X}	1.34	1.61	1.19	6.31	7.25	5.80	7.64	8.86	6.99
SD	1.09	1.10	1.06	2.72	2.59	2.65	3.53	3.38	3.44
n	733	256	477	733	256	477	733	256	477
Ordenador									
\bar{X}	1.51	1.84	1.31	6.52	7.48	5.97	8.03	9.32	7.28
SD	1.07	1.05	1.03	2.43	2.32	2.32	3.20	3.07	3.03
n	634	234	400	634	234	400	634	234	400
d	-.16	-.21	-.11	-.08	-.09	-.07	-.12	-.14	-.09
F	8.69**	5.66*	3.14	2.39	1.1	.93	4.52*	2.53	1.72

Nota. \bar{X} = media; n = tamaño de la muestra; SD = desviación típica; d = d de Cohen; F = F de Fisher; TRC-3 = TRC original (Frederick, 2005); TRC-11 = nuevos 11 ítems incorporados por Salgado (2014a); TRC-14 = ítems Frederick (2005) y 11 ítems Salgado (2014a); N Total. = muestra total; H = hombres; M = mujeres.

* $p < .05$; ** $p < .01$

Análisis Factorial

Con el objetivo de examinar la estructura interna del nuevo TRC, se llevó a cabo un análisis factorial exploratorio (AFE) del test mediante un análisis paralelo. La Tabla 20 recoge la matriz de correlaciones entre los ítems. Las correlaciones presentadas son tetracóricas de acuerdo con lo expuesto sobre la dicotomización de los ítems. Como se puede observar, todos los ítems correlacionan con los demás entre .12 y .73, a excepción del 10 que muestra correlaciones inferiores a .11 con prácticamente la totalidad de los ítems. El estadístico de Bartlett y el test de Kaiser, Meyer y Olkin (KMO) mostraron la adecuación de la matriz de correlaciones para ser factorizada (véase Tabla 21). El estadístico de Bartlett fue significativo y el test KMO mostró un resultado de .868.

Tras el análisis factorial, los resultados reflejaron la existencia de un único factor latente. Aunque tradicionalmente, para seleccionar el número de factores se ha utilizado la regla de Kaiser, esto es, seleccionar los factores con autovalores superiores a 1, la recomendación actual desaconseja esta opción (Lloret-Segura et al., 2014). El principal inconveniente de la regla de Kaiser reside en que el número de factores identificados está directamente relacionado con el número de ítems que analiza (Lloret-Segura et al., 2014). En su lugar, se pueden usar procedimientos como el análisis paralelo (implementado por el programa FACTOR) que identifica los componentes o factores que presentan autovalores superiores a los que se obtendrían por azar. De acuerdo con este análisis, el número de factores identificados en el TRC-14 fue 1. De acuerdo con el método Kaiser, el número de factores identificados serían 3. No obstante, el segundo y tercer factor apenas explican varianza del constructo (proporción de varianza = .092 y .083, respectivamente). Además, el valor de Congruencia Unidimensional (*UniCo*) fue .931, el valor de varianza común explicada (*ECV*) fue .84 y el valor de la media de cargas residuales absolutas del ítem (*MIREAL*) fue .223. Aunque, para que los datos puedan ser tratados como esencialmente unidimensionales se requiere que los Valores de *UniCo* sean superiores a .95, los valores de *ECV* sean inferiores a .30 y los valores de *MIREAL* sean superiores a .85 (Ferrando y Lorenzo-Seva, 2017), los valores obtenidos fueron próximos a ellos.

La Tabla 22 recoge los autovalores (*eigenvalues*) de los factores, así como la varianza explicada del constructo por los factores. Por su parte, la Tabla 23 muestra las cargas factoriales y la comunalidad de los ítems.

Tabla 20.
Correlaciones Tetracóricas entre Ítems

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Ítem 1	-													
Ítem 2	.40***	-												
Ítem 3	.56***	.50***	-											
Ítem 4	.39***	.59***	.55***	-										
Ítem 5	.41***	.50***	.58***	.59***	-									
Ítem 6	.73***	.34***	.59***	.42***	.44***	-								
Ítem 7	.40***	.33***	.45***	.40***	.42***	.42***	-							
Ítem 8	.21***	.15***	.24***	.30***	.17***	.22***	.17***	-						
Ítem 9	.36***	.69***	.47***	.68***	.53***	.39***	.37***	.26***	-					
Ítem 10	.11***	.07***	.10***	.10***	.10***	.11***	.06***	.04***	.06***	-				
Ítem 11	.27***	.28***	.43***	.24***	.35***	.34***	.28***	.15***	.32***	.42***	-			
Ítem 12	.36***	.27***	.41***	.30***	.32***	.41***	.38***	.16***	.31***	.06***	.23***	-		
Ítem 13	.30***	.25***	.39***	.29***	.34***	.32***	.26***	.19***	.30***	-.04***	.24***	.26***	-	
Ítem 14	.38***	.28***	.42***	.22***	.33***	.37***	.29***	.12***	.32***	.10***	.27***	.26***	.22***	-

Nota. * $p < .05$; ** $p < .01$; *** $p < .001$

Tabla 21.

Adecuación de la Matriz de Correlaciones

Determinante de la matriz	.105119772385071
Estadístico de Bartlett	3064.7 ($gl = 91$; $p = .000010$)
Test de Kaiser, Meter y Olkin (KMO)	.86817

Nota. gl = grados de libertad

Tabla 22.

Autovalores y Varianza Explicada Basada en los Autovalores

Variable	Autovalor	Proporción de varianza	Proporción acumulada de varianza
1	5.45035	.38931	.38931
2	1.28087	.09149	
3	1.15689	.08263	
4	.93678	.06691	
5	.81015	.05787	
6	.76788	.05485	

Tabla 23.

Cargas Factoriales y Comunalidad de los Ítems

Variable (ítem)	Factor 1	Comunalidad
1	.672	.452
2	.658	.434
3	.796	.634
4	.713	.508
5	.711	.505
6	.701	.491
7	.575	.331
8	.311	.097
9	.708	.501
10	.157	.025
11	.476	.227
12	.500	.250
13	.448	.201
14	.477	.227

Los índices de ajuste del modelo fueron favorables. Así, el índice *RMSEA* reflejó un valor de .052 y el valor del índice *CFI* fue de .982.

Los resultados de la Tabla 23 muestran que el ítem 10 correlaciona con el factor en .157 por lo que solo puede explicar el 2.5% de la varianza. La convención estándar empleada para asignar los ítems a los factores indica que se deben retener los ítems con saturaciones que estén por encima de .30 o .40 (criterio más conservador; Lloret-Segura et al., 2014). Así pues, el ítem 10 no cumple esta convención. Del mismo modo, el ítem 8 aunque mostró una relación de .311 con el factor, esta fue inferior a .40, por lo que se procedió a realizar dos nuevos análisis factoriales: 1) eliminando el ítem 10 y 2) eliminando los ítems 10 y 8.

Los resultados de los reanálisis se encuentran recogidos en las Tablas 24 y 25. La Tabla 24 muestra los autovalores y la varianza explicada basada en los autovalores. La Tabla 25 recoge las cargas factoriales y la comunalidad de los ítems.

Los datos nuevamente mostraron la existencia de un único factor para ambos reanálisis. El valor de Congruencia Unidimensional (*UniCo*) fue de .978 para el análisis excluyendo los ítems 8 y 10 y de .98 para el análisis excluyendo el ítem 10. Por su parte, el valor de varianza común explicada (*ECV*) fue de .84 para el análisis excluyendo los ítems 8 y 10 y de .843 para el análisis excluyendo el ítem 10. Finalmente, el valor de la media de cargas residuales absolutas del ítem (*MIREAL*) fue de .236 para el análisis excluyendo los ítems 8 y 10 y de .221 para el análisis excluyendo el ítem 10. Teniendo en cuenta que valores de *UniCo* superiores a .95, valores de *ECV* superiores a .85 y valores de *MIREAL* inferiores a .30 sugieren que los datos pueden ser tratados como esencialmente unidimensionales, se sostiene así el ajuste de un único factor.

Tabla 24.

Autovalores y Varianza Explicada Basada en los Autovalores Excluyendo los ítems 8 y 10 o el ítem 10

Variable (ítem)	Análisis excluyendo ítems 8 y 10			Análisis excluyendo ítem 10		
	Autovalor	Proporción de varianza	Proporción acumulada de varianza	Autovalor	Proporción de varianza	Proporción acumulada de varianza
1	5.31882	.44323	.44323	5.42322	.41717	.41717
2	1.18478	.09873		1.18668	.09128	
3	.84019	.07002		.94520	.07271	
4	.80876	.06740		.83937	.06457	
5	.74865	.06239		.78758	.06058	
6	.70065	.05839		.74395	.05723	

Tabla 25.

Cargas Factoriales y Comunalidad de los Ítems Excluyendo los Ítems 8 y 10 o el Ítem 10

Variable (Ítem)	Análisis excluyendo ítems 8 y 10		Análisis excluyendo ítem 10	
	Factor 1	Comunalidad	Factor 1	Comunalidad
1	.672	.452	.672	.452
2	.664	.441	.660	.435
3	.798	.637	.797	.635
4	.708	.501	.714	.509
5	.716	.512	.711	.506
6	.700	.490	.701	.491
7	.577	.333	.576	.332
8	-	-	.312	.097
9	.707	.500	.710	.504
10	-	-	-	-
11	.464	.215	.464	.215
12	.501	.251	.501	.251
13	.449	.201	.452	.204
14	.478	.228	.475	.226

Los resultados de la Tabla 25 muestran nuevamente que la correlación del ítem 8 es inferior a .40 ($r = .312$). No obstante, los índices de ajuste del modelo demostraron un mejor ajuste general del análisis factorial incluyendo el ítem 8 (véase Tabla 26) que excluyéndolo. De este modo, el TRC-14 quedó configurado, finalmente, con 13 ítems (se excluye el ítem 10) y no con 14 problemas como originalmente se había desarrollado. Como consecuencia, los siguientes estudios empíricos de esta tesis doctoral emplearán el TRC desarrollado por Salgado (2014a) excluyendo el ítem 10 (en adelante, TRC-13; véase Apéndice D para el test definitivo y sus estadísticos descriptivos).

Tabla 26.

Índices de Ajuste del Modelo Excluyendo el Ítem 10 o los Ítems 8 y 10

	<i>RMSEA</i>	<i>CFI</i>
Análisis excluyendo ítems 8 y 10	.053	.986
Análisis excluyendo ítem 10	.047	.987

Nota. *RMSEA* = raíz del residuo cuadrático promedio de aproximación; *CFI* = índice de bondad de ajuste comparativo.

Finalmente, la Tabla 27 recoge las correlaciones finales entre el TRC-3, el TRC-10 (anterior TRC-11 sin ítem 10) y el TRC-13 (anterior TRC-14 sin ítem 10) para el total de la muestra, los hombres y las mujeres. En el caso del TRC-3, las correlaciones halladas fueron

correlaciones policóricas. En el resto de los casos, los coeficientes mostrados son correlaciones de Pearson.

Tabla 27.

Correlaciones entre el TRC-3, TRC-10 y TRC-13

	TRC-3 ^a	TRC-10
<i>Muestra Total</i>		
TRC-3	-	
TRC-10	.645	-
TRC-13	.817	.968
<i>Hombres</i>		
TRC-3	-	
TRC-10	.615	-
TRC-13	.809	.961
<i>Mujeres</i>		
TRC-3	-	
TRC-10	.628	-
TRC-13	.806	.966

Nota. Todas las correlaciones son significativas al nivel $p < .001$. TRC-3 = TRC original (Frederick, 2005); TRC-10 = nuevos ítems desarrollados por Salgado (sin ítem 10); TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a).

^a Las correlaciones con el TRC-3 son correlaciones tetracóricas

Fiabilidad del TRC

La Tabla 28 recoge los coeficientes de fiabilidad del TRC-3, TRC-10 y TRC-13 para el total de la muestra, las mujeres y los hombres. El primer coeficiente de fiabilidad calculado fue el coeficiente por consistencia interna Kuder-Richardson 20 ($KR20$). Seguidamente, se calculó el coeficiente de fiabilidad por consistencia interna EAP, calculado mediante el programa FACTOR. El tercer coeficiente de fiabilidad calculado fue el coeficiente test-retest. Entre la realización del TRC en el momento temporal 1 y el momento temporal 2 transcurrieron en promedio 96.67 días ($SD = 170.90$) y la muestra empleada para el cálculo de este coeficiente se compuso de 295 sujetos. Para el cálculo del coeficiente de fiabilidad test-retest del TRC-3 se emplearon las correlaciones policóricas obtenidas mediante el programa FACTOR, por tratarse de una variable discreta con 4 categorías de puntuaciones (0, 1, 2 o 3) y tener una distribución sesgada. Por último, se calculó el coeficiente de fiabilidad *CES*. Para estimar este coeficiente, en el caso del TRC-10 se empleó la fórmula 14 propuesta por Schmidt et al. (2003) y para el caso del TRC-13 se empleó la fórmula 17.

Tabla 28.
Coefficientes de Fiabilidad de los Test de RC

	TRC-3			TRC-10			TRC-13		
	N Total	Hombres	Mujeres	N Total	Hombres	Mujeres	N Total	Hombres	Mujeres
Fiabilidad <i>KR20</i>	.58	.58	.55	.70	.68	.67	.78	.77	.76
Fiabilidad EAP	.78	.75	.80	.85	.88	.83	.90	.91	.89
Fiabilidad Test-retest	.71	.76	.69	.72	.74	.68	.78	.82	.74
Fiabilidad <i>CES</i>	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>	.57	.52	.53	.66	.67	.61

Nota. TRC-3 = TRC original (Frederick, 2005); TRC-10 = nuevos ítems incorporados por Salgado (2014a) eliminando el ítem 10; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); *n.a.* = no aplica; *KR20* = Kuder-Richardson 20; *CES* = coeficiente de equivalencia y estabilidad.

Esta última es una variante de la fórmula 14 y se aplica en los casos en que el test está compuesto por un número impar de ítems y, por ende, no es posible dividirlo en dos mitades paralelas. En el caso del TRC-3 no se pudo calcular el coeficiente *CES* por estar compuesto, únicamente, por 3 ítems.

Como se puede observar en la Tabla 28, las magnitudes de los coeficientes de fiabilidad encontrados fueron aceptables, salvo los indicados para el TRC-3. En general, la magnitud de la fiabilidad de los test es mayor para el TRC-13, con independencia de la muestra empleada y del coeficiente de fiabilidad calculado. Respecto al tipo, los coeficientes de fiabilidad EAP calculados mediante el programa FACTOR muestran magnitudes más robustas que los encontrados mediante la *KR20*, estabilidad temporal o *CES*. No obstante, estos coeficientes no son comparables entre sí pues, cada uno de ellos estima el error de medida producido por fuentes de error diferentes. En cualquier caso, el coeficiente *CES* es el que calcula la magnitud del error de medida controlando más fuentes de error. Por este motivo, la estimación de este coeficiente siempre debería ser inferior a la de otros coeficientes de fiabilidad. Sin embargo, es el método más apropiado para estimar la fiabilidad de una medida (Salgado, 2015; Schmidt et al., 2003), ya que al controlar más fuentes de error no sobreestima la magnitud de la fiabilidad de la medida, y por consiguiente no subestima la correlación verdadera (ρ) entre constructos (Schmidt y Hunter, 1996).

En resumen, los datos de la Tabla 28 indican que el TRC-13 mide la RC con menor grado de error que el TRC-3, con independencia del coeficiente de fiabilidad empleado, confirmandose, así, la Hipótesis 4.

4. DISCUSIÓN

Diversos estudios han puesto de manifiesto la necesidad de crear nuevos test de RC (Salgado et al., 2019; Toplak et al., 2014; Weller et al., 2013). Esta necesidad no solo se ha justificado por la gran popularidad y visibilidad de los 3 ítems del test de Frederick (2005; Kahneman y Frederick, 2002) sino también por las evidentes limitaciones psicométricas del mismo. El principal objetivo de este estudio ha sido explorar las propiedades psicométricas de un nuevo test de RC (Salgado, 2014a) y comprobar si este supera las limitaciones enumeradas en la literatura científica para el TRC original.

Los resultados han reflejado que el nuevo TRC es superior al TRC-3, desde el punto de vista psicométrico, en varios aspectos. En primer lugar, el TRC-13 muestra una distribución de las puntuaciones aproximadamente normal con valores de asimetría y curtosis inferiores a 1 en valores absolutos. Sin embargo, el TRC-3 refleja una distribución sesgada con forma binomial y valores de curtosis superiores a 1. En segundo lugar, este estudio ha mostrado que el grado de dificultad del TRC-3 y del TRC-13 es diferente. Mientras que las puntuaciones del TRC-3 se agrupan principalmente en los valores bajos de la variable (asimetría positiva) en el TRC-13, las puntuaciones se agrupan mayoritariamente en los valores medio-altos de la variable (asimetría

negativa). Los ítems con un menor porcentaje de aciertos han sido el 1, 3, 8, 11, 13 y 14, donde menos del 50% de la muestra ha conseguido responder de forma correcta.

En tercer lugar, el análisis de fiabilidad de los tests ha indicado que el TRC-13 y el TRC-10 presentan unos coeficientes de fiabilidad más altos que el TRC-3, con independencia del tipo de coeficiente estimado. Estudios previos apenas han examinado la fiabilidad del test de RC y, en aquellos casos en que sí lo han hecho, han calculado el coeficiente de consistencia interna *KR20*. Los análisis estadísticos llevados a cabo para estimar la fiabilidad por consistencia interna de un test asumen una distribución normal de las puntuaciones. Sin embargo, en el caso del TRC-3, esta propiedad no se cumple, lo que provoca que las estimaciones efectuadas al respecto no sean necesariamente robustas. Además, este coeficiente controla dos fuentes de error muy concretas. No obstante, otras fuentes de error en la medida no han sido tenidas en cuenta en el cálculo de la fiabilidad del test. En este sentido, el estudio hace una importante contribución a la literatura científica de la RC estimando la fiabilidad del test mediante cuatro coeficientes diferentes: coeficiente *KR20*, coeficiente calculado mediante las puntuaciones de la distribución posterior (EAP), coeficiente test-retest y el coeficiente *CES*, *no examinados hasta el momento*. Cada uno de ellos controla diferentes tipos de error. Únicamente el coeficiente *CES* estima la fiabilidad de la medida controlando 3 fuentes de error (errores por respuestas aleatoria, errores específicos de factor y errores por el paso del tiempo), configurándose como el mejor estimador de la fiabilidad real del test. Esta investigación presenta la primera estimación del coeficiente *CES* del TRC, siendo una contribución única de este trabajo.

Otro de los objetivos de esta investigación ha sido explorar la distribución de las puntuaciones del TRC en hombres y mujeres. Los resultados encontrados en este estudio están en línea con los encontrados en investigaciones previas (Brañas-Garza et al., 2015; Primi et al., 2018). Así, las mujeres de esta muestra han mostrado un peor desempeño que los hombres en los tres instrumentos de RC. Estas diferencias han sido significativas en todos los casos y parecen aumentar a medida que se incrementa el número de ítems del test. Asimismo, se han encontrado diferencias significativas en función del sexo en cuanto a la proporción de sujetos que aciertan cada ítem. Los resultados indican que, tanto para la muestra de hombres como para la muestra de mujeres, el TRC-13 refleja una distribución de las puntuaciones aproximadamente normal, con valores de asimetría y curtosis inferiores a 1 en valores absolutos y, en algunos casos, próximos a 0; mientras que el TRC-3 refleja una distribución sesgada, con valores de curtosis negativos superiores a 1 para los hombres y positivos para las mujeres. En el TRC-10, la distribución de las mujeres es prácticamente simétrica, mientras que la de los hombres es asimétrica negativa.

Los datos, en su conjunto, reflejan, una vez más, las diferencias individuales en función del sexo en la RC. Diversos estudios han tratado de justificar la presencia de estas diferencias. Así, algunos autores han sugerido que las diferencias en función del sexo podrían ser debidas a las diferencias individuales en la capacidad cognitiva o capacidad matemática (Morsanyi et al.,

2014; Primi et al., 2015; Primi et al., 2018; Thomson y Oppenheimer, 2016) y otros han indicado que, el ratio 2D:4D puede explicar parcialmente las diferencias de sexo en el desempeño en el TRC (Bosch-Domènech et al., 2014; Cueva et al., 2015). Del mismo modo, Campitelli y Gerrans (2014) encontraron que variables muy similares, aunque diferentes, predicen el desempeño en el TRC en hombres y mujeres. En este sentido, este estudio ha mostrado que un posible determinante del desempeño de los hombres en el TRC-3 es el formato de administración del TRC (papel y lápiz vs. ordenador). De modo que los hombres que han respondido al test en el ordenador han conseguido un mejor desempeño que los que han contestado mediante lápiz y papel, aunque estas diferencias fueron pequeñas y únicamente se encontraron en la evaluación de la RC con el TRC-3. Es necesario llevar a cabo nuevos estudios que nos permitan profundizar en el conocimiento de las diferencias en función del sexo en la RC.

Otra de las propiedades psicométricas examinadas ha sido la estructura factorial del test. Los resultados han indicado que los ítems del nuevo test cargaron en un único factor, al igual que previos estudios han sugerido para el TRC-3 (Primi et al., 2015). El modelo de mejor ajuste resultó ser aquel en el que el TRC se conforma por 13 ítems (eliminando el ítem 10) y esta versión se utilizará en los posteriores estudios de esta tesis.

Por último, el estudio presenta algunas limitaciones. En primer lugar, la muestra empleada en la investigación es una muestra conformada únicamente por estudiantes universitarios. Aunque el tamaño de la muestra es grande y se podría considerar representativo de la población, se recomendaría replicar este estudio con otras muestras diferentes. Además, como el TRC es un test muy vinculado a las capacidades cognitivas, muy probablemente exista restricción en el rango en nuestra muestra, lo que sería necesario neutralizar para hacer una mejor estimación de la media y desviación típica real del TRC en las tres muestras (muestra total, hombres y mujeres).

Por otro lado, el cálculo del coeficiente de fiabilidad test-retest ha sido hallado con una muestra de 295 sujetos. Aunque su tamaño es importante, no alcanza la cifra considerada como tamaño de la muestra grande ($N = 400$), por lo que, debería ampliarse en posteriores estudios.

Para resumir, esta investigación nos ha permitido examinar las propiedades psicométricas de un nuevo y más amplio TRC (Salgado, 2014a) y compararlas con el TRC original (Frederick, 2005; Kahneman y Frederick, 2002). De acuerdo con las hipótesis de partida, los resultados han mostrado que el nuevo test es superior al original, desde el punto de vista psicométrico, al eliminar las limitaciones del TRC original, mejorar la fiabilidad del test y permitir que las puntuaciones se distribuyan de forma normal o aproximadamente normal. Asimismo, este estudio ha contribuido al conocimiento de la RC al examinar la magnitud de cuatro tipos diferentes de coeficientes de fiabilidad y ha contribuido al conocimiento de la estructura factorial del instrumento.



ESTUDIO 2.

Efectos del Orden de Presentación de los Ítems sobre el Resultado en el TRC

1. INTRODUCCIÓN

El objetivo de este estudio fue comprobar los efectos que determinados factores relacionados con la aplicación y administración del test pueden tener sobre el desempeño en el nuevo TRC. Se examinaron 3 cuestiones: (1) el efecto que el orden de presentación de los ítems tiene sobre el resultado en el TRC-13, (2) los efectos que el orden de presentación de los ítems tiene sobre el tiempo de respuesta al test, (3) la relación entre el tiempo de respuesta y el desempeño en el TRC-13 y (4) los efectos y relaciones anteriores en el grupo de hombres y de mujeres de forma independiente.

Investigaciones previas han sugerido que la secuencia de presentación de los ítems del TRC-3 puede afectar al resultado final de los participantes. Concretamente, la integración de Brañas-Garza et al. (2015) mostró que los sujetos obtuvieron mejores resultados en el TRC cuando la secuencia de ítems se presentó en el orden original (bate y pelota, máquinas y nenúfares) que cuando se utilizó cualquiera de las otras posibles combinaciones. El primer objetivo de este estudio ha sido examinar los efectos que el orden de presentación de los ítems tiene sobre los resultados de un nuevo TRC. A este respecto, la hipótesis planteada es:

Hipótesis 1. *El orden de presentación de los ítems afecta a los resultados en el TRC.*

Una de las cuestiones más ampliamente examinadas cuando se pretenden mostrar evidencias de validez basadas en los procesos de respuesta es la relacionada con el tiempo de reacción. En relación con esto, estudios previos han examinado la fluidez de respuesta en el TRC para comprobar si el tiempo invertido en responder a los ítems podría ser un indicativo del modo de procesamiento cognitivo empleado en la realización del test (procesamiento de tipo 1 vs. procesamiento de tipo 2). Los principales resultados han señalado que es más probable que las respuestas rápidas reflejen resultados heurísticos o basados en un procesamiento de tipo 1 que las

respuestas lentas. Por lo mismo, puntuaciones altas en el TRC tienden a estar asociadas con la inversión de más tiempo de reacción (Anderson et al., 2008; Böckenholt, 2012; Szaszi et al., 2017; Thompson et al., 2011; Travers et al., 2016). El segundo objetivo de este estudio ha sido examinar si el orden de presentación de los ítems afecta al tiempo de respuesta a los mismos. No obstante, como esta relación no ha sido testada hasta el momento, no se proponen hipótesis al respecto.

Asimismo, el tercer objetivo de este estudio ha sido determinar la relación entre el resultado en el TRC y el tiempo invertido en responder al test. En este sentido la hipótesis planteada es:

Hipótesis 2: *El tiempo invertido en responder al TRC se relaciona positivamente con los resultados en el test.*

Finalmente, las evidencias basadas en el proceso de respuesta también pueden contribuir a explicar las diferencias en el significado o interpretación de las puntuaciones del test entre subgrupos examinados. En el estudio 1 de esta tesis se han encontrado diferencias en función del sexo en el desempeño en el TRC. Aunque se han investigado diversas variables que pueden actuar como determinantes de estas diferencias (p. ej. capacidad matemática, ansiedad matemática, ratio 2D:4D, etc.), apenas hay estudios que examinen los procesos de respuesta para ambos grupos. El tercer objetivo de esta investigación ha sido examinar si el orden de presentación de los ítems determina el desempeño de hombres y mujeres y si afecta al tiempo que ambos grupos invierten en responder al TRC. Asimismo, se examinan las posibles diferencias existentes entre ambos grupos en cuanto al desempeño y al tiempo invertido en responder al test, según el orden de presentación de los ítems. En ambos casos, no se proponen hipótesis al respecto.

2. MÉTODO

Muestra

La muestra inicial recogida para este estudio se compuso de 634 sujetos. Del total de la muestra, un 5.05% ($n = 32$) fueron eliminados por ser considerados *outliers* en cuanto al tiempo invertido en la realización del TRC. Por lo tanto, la muestra final empleada se compuso de 602 sujetos, de los cuales el 63% fueron mujeres ($n = 379$). La edad promedio de los sujetos fue 21.39 años ($SD = 3.83$, rango = 18 a 50 años). Todos los participantes eran estudiantes universitarios y el 2.82% de ellos eran postgraduados. El 97.18% restante de la muestra estaba compuesto por alumnos matriculados en diversas titulaciones. La Tabla 29 recoge las frecuencias de los estudiantes según el área de estudio al que pertenece su titulación. Como se puede observar, el 58.97% de los participantes eran alumnos matriculados en alguna titulación de grado del área de las ciencias sociales, el 11.11% de los participantes pertenecen a alguna titulación de ciencias

experimentales o tecnológicas, el 21.88% cursaron alguna titulación de ciencias de la salud y el 6.67% del área de humanidades.

Tabla 29.

Características de la Muestra (N = 602)

	<i>n</i>	%
Sexo		
Mujeres	379	63.00
Hombres	223	37.00
Edad		
≤ 19	178	29.57
20-22	292	48.50
> 22	132	21.93
Área de Estudio		
Ciencias sociales	345	58.97
Ciencias experimentales	65	11.11
Ciencias de la salud	128	21.88
Humanidades	39	6.67
No indicado	8	1.37

Procedimiento

Para llevar a cabo este estudio experimental, se solicitó la participación voluntaria de estudiantes universitarios mediante la publicación de anuncios colgados en facultades y otros espacios públicos de la Universidad de Santiago de Compostela (p. ej., instalaciones deportivas, bibliotecas, unidades de gestión académica o residencias universitarias). La recogida de muestra se llevó a cabo en dos fases: la primera, entre los meses de enero y junio de 2017; la segunda, entre esos mismos meses del año siguiente, 2018. El 63.6% de la muestra ($n = 383$) fue reclutada en el 2017 y los restantes en el 2018 ($n = 220$). Con el objetivo de atraer a los estudiantes y maximizar su desempeño en el estudio, se les ofreció una compensación económica a cambio de su participación. El 50% de los estudiantes que participaron en el 2017 recibieron 10€ en contraprestación y participaron en un sorteo de 100€. El 50% restante no fueron incentivados económicamente a cambio de su participación, sino que se les entregó un informe con los resultados obtenidos en las diversas pruebas administradas. Sin embargo, el 97% de los sujetos que participaron en el 2018 recibieron la contraprestación económica de 15€ y solo un 3% recibieron el informe con sus resultados. Para hacer funcional el experimento, se organizaron 62 grupos de entre 10 y 15 personas.

Durante el desarrollo de las sesiones se instruyó a los participantes en el carácter confidencial y anónimo del estudio y en el único propósito e interés por examinar los resultados

de modo agregado y con fines de investigación. Asimismo, antes de iniciar el estudio, los participantes firmaron un consentimiento informado sobre el mismo.

Con el objetivo de calcular la fiabilidad test-retest de las medidas empleadas en este estudio se animó a los sujetos a participar de nuevo en el estudio experimental. Entre la realización del TRC en el momento temporal 1 y 2 transcurrieron, en promedio, 96.67 días ($SD = 170.90$). En este segundo caso, la muestra se redujo a 295 sujetos, de los cuales 210 fueron mujeres. La edad promedio de la muestra fue de 21.89 ($SD = 4.12$) y el rango de edad abarcó desde los 17 a los 58 años.

Preparación del Experimento e Instrumentos de Medida

Todos los participantes completaron el TRC-13 (Salgado, 2014a; véase Apéndice D) además de otras pruebas con un propósito diferente al de este estudio. El TRC fue presentado en formato ordenador con el objetivo de poder registrar el tiempo de respuesta de los ítems. Para ello, se empleó el software de Millisecond Inquisit 5 Lab (Inquisit, 2016).

Cada uno de los ítems que conforman el test fue presentado de forma individual en el centro de la pantalla del ordenador. Para responder, los participantes tuvieron que seleccionar una de las dos opciones de respuesta propuestas de forma paralela e inmediatamente debajo del enunciado del ítem. Una vez completado el ítem, se debía presionar el botón “siguiente” situado en el margen inferior derecho de la pantalla para continuar con el test. Para pasar al posterior ítem, el programa requería haber respondido a la pregunta previa. Una vez presionado el botón “siguiente”, el programa informático no permitía volver a las preguntas anteriores ni modificar las respuestas. Los participantes fueron informados de estas particularidades antes de ser administrado el test, pero en ningún caso se informó de que se registrarían los tiempos de respuesta para evitar cualquier efecto que esta información pudiese tener sobre el desempeño en la tarea.

Los estudiantes no dispusieron de límite temporal para completar el test. Por cada ítem respondido correctamente, recibieron un punto y por cada ítem respondido erróneamente 0 puntos, de modo que el rango de puntuaciones osciló entre 0 y 13.

Durante el desarrollo del estudio se manipuló el orden secuencial de presentación de los ítems, de modo que se diseñaron tres condiciones experimentales: (1) se presentaron los 3 ítems de Frederick (2005; Kahneman y Frederick, 2002) seguidos de los 10 nuevos ítems propuestos por Salgado (2014a); (2) se presentaron los ítems 4 a 9 de Salgado (2014a) seguidos de los 3 ítems originales de Frederick (ítems 1 a 3) y de los ítems 10 a 13 de Salgado; (3) se presentaron los ítems 4 a 13 de Salgado seguidos de los 3 ítems de Frederick. De este modo, las diversas condiciones se configuraron en función de la posición del TRC-3 dentro del TRC-13: al inicio,

en medio o al final, para la condición 1, 2 y 3, respectivamente. Los estudiantes fueron asignados de forma aleatoria a una de estas tres condiciones. La Tabla 30 recoge las frecuencias y porcentajes de sujetos que participaron en cada una de las condiciones experimentales.

Tabla 30.

Distribución de la Muestra por Condiciones Experimentales

	Condición 1	Condición 2	Condición 3	Total
Hombres	69 (35%)	68 (34.2%)	86 (41.7%)	223 (37%)
Mujeres	128 (65%)	131 (65.8%)	120 (58.3%)	379 (63%)
Total	197 (32.7%)	199 (33.1%)	206 (34.2%)	602 (100%)

3. RESULTADOS

Efectos del Orden de Presentación de los Ítems sobre los Resultados en el TRC

La Tabla 31 refleja los efectos que el orden de presentación de los ítems tiene sobre los resultados en el TRC. La descripción de los resultados se lleva a cabo mostrando los efectos que el orden de presentación de los ítems tiene sobre el desempeño en cada uno de los ítems en el TRC-13, así como en el sumatorio del total de los ítems (TRC-13), de los nuevos ítems (TRC-10) y de los ítems originales (TRC-3). Asimismo, se indican las diferencias entre condiciones experimentales en el desempeño del test, expresadas con el estadístico F de Fisher y el tamaño del efecto d de Cohen (Cohen y Cohen, 1983).

Como se puede observar, los resultados muestran que, cuando se sitúan los ítems originales en medio de los nuevos ítems (condición 2), los participantes alcanzan significativamente mejores resultados en el TRC-13 y en los ítems originales (TRC-3), que cuando se sitúan en cualquiera de las otras dos combinaciones (al inicio, $d = -.21$ y $d = -.47$, respectivamente, o al final del test, $d = .23$ y $d = .27$, respectivamente). Los datos también reflejaron que los tamaños del efecto fueron mayores para los ítems del TRC-3 que para los restantes ítems. De modo que los resultados en los ítems originales se ven más afectados por el orden de presentación de los ítems.

Asimismo, los datos también han indicado que, cuando se sitúan los ítems originales al final de la secuencia total de preguntas (condición 3), los participantes alcanzan significativamente mejores resultados en los ítems originales (TRC-3) que cuando se sitúan al inicio de la secuencia ($d = -.22$). No obstante, el rendimiento en los restantes ítems del TRC (Salgado, 2014a), no se ve afectado por la ubicación de los ítems originales a lo largo del test.

Tabla 31.

Medias, Desviaciones Típicas y Diferencias en el Desempeño en el TRC según la Condición Experimental (Muestra Total)

Desempeño en el test									Diferencias en el desempeño								
Condición 1			Condición 2			Condición 3			Condición 1 vs. 2			Condición 1 vs. 3			Condición 2 vs. 3		
\bar{X}	<i>SD</i>		\bar{X}	<i>SD</i>		\bar{X}	<i>SD</i>		<i>d</i>	<i>F</i>		<i>d</i>	<i>F</i>		<i>d</i>	<i>F</i>	
Ítem 1	.32	.47	.52	.50		.46	.50		-.41	17.35***		-.29	8.58**		.12	1.53	
Ítem 2	.51	.50	.70	.46		.63	.48		-.40	15.59***		-.25	6.33*		.15	2.06	
Ítem 3	.45	.50	.55	.50		.41	.49		-.20	4.06*		.08	.62		.28	8.07**	
Ítem 4	.75	.44	.72	.45		.68	.47		.07	.26		.15	1.88		.09	.74	
Ítem 5	.56	.50	.61	.49		.60	.49		-.10	1.00		-.08	.47		.02	.11	
Ítem 6	.66	.48	.70	.46		.70	.46		-.09	.87		-.09	.71		.00	.01	
Ítem 7	.53	.50	.62	.49		.48	.50		-.18	2.94 [†]		.10	1.10		.28	7.84**	
Ítem 8	.48	.50	.44	.50		.42	.50		.08	.81		.12	1.46		.04	.09	
Ítem 9	.81	.40	.85	.36		.76	.43		-.11	1.23		.12	1.20		.23	4.93*	
Ítem 10	.46	.50	.51	.50		.45	.50		-.10	.82		.02	.04		.12	1.27	
Ítem 11	.72	.45	.80	.40		.72	.45		-.19	3.33 [†]		.00	.00		.19	3.60 [†]	
Ítem 12	.47	.50	.44	.50		.45	.50		.06	.24		.04	.10		-.02	.04	
Ítem 13	.50	.50	.44	.50		.42	.50		.12	1.70		.16	2.61		.04	.09	
TRC-13	7.22	3.39	7.90	3.01		7.19	3.11		-.21	4.48*		.01	.01		.23	5.45*	
TRC-3	1.27	1.12	1.77	1.02		1.50	1.01		-.47	21.10***		-.22	4.53*		.27	7.07**	
TRC-10	5.94	2.55	6.13	2.25		5.69	2.41		-.08	.60		.10	1.07		.19	3.64 [†]	

Nota. Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; *SD* = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = sumatorio de los ítems originales (Frederick, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el ítem 10.

† $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Estos datos parecen sugerir que presentar los ítems del TRC-3 en medio de los nuevos ítems podría ayudar a los participantes a alcanzar significativamente mejores resultados en el TRC-13 y a conseguir que más gente responda correctamente a los ítems originales. En su conjunto estos resultados apoyan la Hipótesis 1.

Otro de los propósitos de este estudio ha sido examinar los efectos que el orden de presentación de los ítems tiene sobre el resultado en el TRC para la muestra de hombres y de mujeres de forma independiente. Para operativizar y facilitar la lectura del estudio, a continuación, se resume los resultados encontrados al respecto y se relega al Apéndice E y F las tablas con los resultados obtenidos en ambos grupos.

Los datos muestran que los resultados encontrados para la muestra de hombres y de mujeres son similares a los encontrados para la muestra total. Así pues, cuando se sitúan los ítems originales en medio de los nuevos ítems (condición 2), tanto los hombres como las mujeres alcanzan significativamente mejores resultados en los ítems originales (TRC-3) que cuando se sitúan en cualquiera de las otras dos combinaciones (al inicio, $d_H = -.54$ y $d_M = -.44$, respectivamente; o al final del test, $d_H = .34$ y $d_M = .29$, respectivamente). En ese mismo caso, las mujeres, aunque no los hombres, también alcanzan una mejora significativa en el desempeño total en el TRC-13 ($d = -.22$ en la condición 1 vs. 2 y $d = .30$ en la condición 2 vs. 3). Finalmente, el rendimiento de las mujeres en los nuevos ítems (TRC-10) también se incrementa significativamente cuando los ítems originales se sitúan en medio de la secuencia que cuando se disponen al final de la misma ($d = .26$). Sin embargo, el desempeño de los hombres en los nuevos ítems (TRC-10) y en el TRC-13 no varía significativamente en función de la posición de los ítems.

Los resultados, en su conjunto, parecen sugerir que tanto en la muestra total como en los grupos de hombres y de mujeres, el orden de disposición de los ítems afecta a los resultados alcanzados en los ítems originales y en la totalidad del TRC-13. Sin embargo, los nuevos ítems parecen controlar mejor los efectos que produce el orden de presentación de los ítems en el desempeño en el test.

Diferencias en Función del Sexo en el Desempeño en el TRC por Condiciones Experimentales

La Tabla 32 refleja las diferencias en función del sexo en el desempeño en el TRC para cada una de las condiciones experimentales. Como se puede observar, los hombres alcanzaron significativamente mejores resultados que las mujeres en la realización de los ítems de los TRC-13, TRC-3 y TRC-10 en las tres condiciones experimentales. Las diferencias fueron moderadas en todos los casos (Cohen, 1977) y los tamaños del efecto oscilaron entre .41 y .61 en la condición 1, entre .55 y .71 en la condición 2 y entre .51 y .70 en la condición 3. Asimismo, los datos reflejan

que las diferencias fueron ligeramente superiores en el TRC-13 que en el TRC-3 o TRC-10, sugiriendo que el número de ítems del test podría moderar las diferencias en función del sexo en el test.

Finalmente, aunque los datos muestran que las diferencias fueron mayores para todos los TRC en la condición 2 que en la condición 1 y 3, las diferencias entre los tamaños del efecto no fueron significativas (rango de $z = .05$ a $.54$); lo que podría indicar que el orden de presentación de los ítems no produce un efecto sobre las diferencias entre hombres y mujeres en el rendimiento en el test.

Efectos del Orden de Presentación de los Ítems sobre el Tiempo de Respuesta (TR) al TRC

La Tabla 33 refleja los efectos que el orden de presentación de los ítems tiene sobre el tiempo de respuesta al TRC. Los datos de la tabla recogen el tiempo medio invertido en responder a cada uno de los ítems, así como al sumatorio del total de los ítems (TRC-13), de los ítems originales (TRC-3) y de los nuevos ítems (TRC-10), en cada condición experimental. Asimismo, se indican las diferencias entre las condiciones en el tiempo de respuesta, expresadas con el estadístico F de Fisher y los tamaños del efecto d de Cohen (Cohen y Cohen, 1983).

Como se puede observar, los resultados muestran que no existen diferencias significativas entre las condiciones experimentales en cuanto al tiempo que los participantes, en promedio, han invertido en responder a los ítems del TRC. Únicamente, en el caso de los nuevos ítems (TRC-10), los participantes invirtieron significativamente más tiempo en responder, cuando los ítems originales se posicionaron al final del test (condición 3) que cuando se dispusieron al inicio (condición 1; $d = -.22$). Del mismo modo, se encontró que los participantes invirtieron marginalmente más tiempo en responder al TRC-3 cuando estos ítems se colocaron al inicio del test (condición 1) que cuando se dispusieron en el medio (condición 2; $d = .19$). Pero en ambos casos, las diferencias halladas fueron pequeñas (Cohen, 1977).

En conjunto, estos datos indican que, aunque el orden de presentación de los ítems no afecta al tiempo invertido en responder al TRC-13, los sujetos tienden a invertir significativamente más tiempo en la deliberación de la respuesta de los ítems que se encuentran posicionados al inicio de la secuencia de ítems.

Tabla 32.

Diferencias entre Hombres y Mujeres en el Desempeño en el TRC según la Condición Experimental

Condición 1						Condición 2						Condición 3								
Hombres			Mujeres			Hombres			Mujeres			Hombres			Mujeres					
\bar{X}	SD		\bar{X}	SD	d	F	\bar{X}	SD		\bar{X}	SD	d	F	\bar{X}	SD		\bar{X}	SD	d	F
Ítem 1	.45	.50	.25	.44	.43	8.45**	.66	.48	.45	.50	.43	8.27**	.52	.50	.42	.50	.20	2.29		
Ítem 2	.57	.50	.48	.50	.18	1.41	.75	.44	.67	.47	.17	1.30	.72	.45	.57	.50	.31	5.20*		
Ítem 3	.55	.50	.39	.49	.32	4.71*	.72	.45	.46	.50	.54	13.16***	.55	.50	.31	.46	.50	12.36***		
Ítem 4	.83	.38	.70	.46	.30	3.61 ¹	.76	.43	.70	.46	.13	.87	.81	.39	.59	.49	.49	12.02***		
Ítem 5	.68	.47	.50	.50	.37	6.11*	.71	.46	.56	.50	.31	3.79 ¹	.67	.47	.54	.50	.27	3.70 ¹		
Ítem 6	.78	.42	.59	.49	.41	7.32**	.85	.36	.63	.49	.49	11.60***	.79	.41	.63	.48	.35	6.01*		
Ítem 7	.67	.48	.46	.50	.43	7.85**	.74	.44	.56	.50	.38	6.14*	.62	.49	.38	.49	.49	11.39***		
Ítem 8	.55	.50	.45	.50	.20	2.00	.51	.50	.40	.49	.22	2.53	.50	.50	.37	.48	.27	3.68 ¹		
Ítem 9	.83	.38	.80	.40	.08	.24	.87	.34	.84	.37	.08	.27	.83	.38	.72	.45	.26	3.30 ¹		
Ítem 10	.54	.50	.42	.50	.24	2.36	.60	.49	.46	.50	.28	3.80 ¹	.55	.50	.38	.49	.34	5.48*		
Ítem 11	.88	.32	.63	.48	.58	15.00***	.96	.21	.72	.45	.62	17.03***	.83	.38	.64	.48	.43	8.65**		
Ítem 12	.57	.50	.41	.49	.32	4.16*	.57	.50	.37	.49	.41	7.42**	.50	.50	.42	.50	.16	1.40		
Ítem 13	.59	.50	.45	.50	.28	3.60 ¹	.53	.50	.39	.49	.28	3.60 ¹	.50	.50	.37	.48	.27	3.68 ¹		
TRC-13	8.48	3.34	6.54	3.22	.60	15.81***	9.24	2.73	7.21	2.93	.71	22.53***	8.38	2.90	6.33	2.98	.70	24.32***		
TRC-3	1.57	1.11	1.12	1.10	.41	7.43**	2.13	.96	1.58	1.01	.55	13.88***	1.79	1.02	1.29	.96	.51	12.92***		
TRC-10	6.91	2.48	5.42	2.44	.61	16.61***	7.10	2.02	5.63	2.20	.69	21.35***	6.59	2.21	5.04	2.34	.68	23.08***		

Nota. Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; *SD* = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = sumatorio de los ítems originales (Frederick, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el 10.

¹ $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Tabla 33.

Medias y Desviaciones Típicas de los Tiempos de Respuesta (en Segundos) del TRC según la Condición Experimental (Muestra Total)

Tiempo de Respuesta (TR)									Diferencias en el TR								
Condición 1			Condición 2			Condición 3			Condición 1 vs. 2			Condición 1 vs. 3			Condición 2 vs. 3		
\bar{X}	SD		\bar{X}	SD		\bar{X}	SD		d	F	d	F	d	F	d	F	
Ítem 1	31.65	26.33	28.66	20.43		32.84	27.28		.13	1.60	-.04	.20	-.17		3.04 ¹		
Ítem 2	39.57	30.25	32.05	27.10		29.87	25.99		.26	6.79**	.35	11.96**	.08		.68		
Ítem 3	40.33	25.89	40.63	20.09		40.21	27.95		-.01	.02	.00	.00	.02		.03		
Ítem 4	31.80	29.61	48.35	52.44		45.99	36.90		-.39	14.92***	-.42	18.03***	.05		.28		
Ítem 5	50.98	39.53	49.98	34.68		56.49	46.19		.03	.07	-.13	1.65	-.16		2.56		
Ítem 6	52.18	41.23	66.00	62.76		64.55	47.04		-.26	6.69**	-.28	7.85**	.03		.07		
Ítem 7	41.71	31.59	39.88	22.70		46.41	31.54		.07	.44	-.15	2.23	-.24		5.68*		
Ítem 8	61.01	53.17	59.75	45.90		64.37	45.19		.03	.06	-.07	.47	-.10		1.04		
Ítem 9	35.37	21.73	37.91	20.61		42.80	27.89		-.12	1.42	-.30	8.84**	-.20		4.01*		
Ítem 10	67.97	50.57	64.95	45.31		71.89	56.16		.06	.39	-.07	.54	-.14		1.86		
Ítem 11	65.55	50.05	63.71	44.25		67.20	45.15		.04	.15	-.04	.12	-.08		.61		
Ítem 12	34.85	28.45	38.73	30.85		43.49	34.30		-.13	1.70	-.27	7.54**	-.15		2.15		
Ítem 13	93.02	77.50	91.45	79.02		87.68	66.68		.02	.04	.07	.55	.05		.27		
TRC-13	646.01	298.93	662.06	272.64		693.79	294.55		-.06	.31	-.16	2.61	-.11		1.26		
TRC-3	111.56	59.32	101.33	45.70		102.92	58.31		.19	3.70 ¹	.15	2.17	-.03		.09		
TRC-10	534.45	261.98	560.72	244.10		590.87	254.50		-.10	1.07	-.22	4.81*	-.12		1.48		

Nota. Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; SD = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = sumatorio de los ítems originales (Frederick, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el 10.

1 $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Por otro lado, otro de los propósitos de este estudio ha sido examinar los efectos que el orden de presentación de los ítems tiene sobre el tiempo de respuesta al test en la muestra de hombres y mujeres. Nuevamente, para operativizar y facilitar la lectura del estudio, a continuación, se explican los resultados encontrados al respecto y para una revisión pormenorizada de los datos véase los Apéndices G y H.

Respecto a las mujeres, los resultados mostraron que no hay diferencias significativas entre las diferentes condiciones experimentales en cuanto al tiempo que las mujeres invierten en responder a los ítems del test (rango de $d = .01$ a $.16$). Sin embargo, en el caso de los hombres, los resultados han replicado los hallados en la muestra total. De modo que los hombres invirtieron significativamente más tiempo en responder a los ítems originales (TRC-3) cuando estos se dispusieron al inicio de la secuencia de ítems (condición 1) que cuando se dispusieron en el medio (condición 2; $d = .36$). Asimismo, invirtieron marginalmente más tiempo en la deliberación de los nuevos ítems (TRC-10), cuando los ítems originales se colocaron al final de la secuencia (condición 3) que cuando se dispusieron al inicio (condición 1; $d = -.29$).

En su conjunto, los resultados indican que el modo de ordenar los ítems del TRC-13 no afecta al tiempo que las mujeres invierten en responder al test ni a los ítems originales y nuevos por separado. Sin embargo, los datos muestran que los hombres tienden a invertir significativamente más tiempo en responder a los ítems que se encuentran posicionados al inicio de la secuencia, aunque esto no incrementa significativamente el tiempo que los hombres invierten en responder a la totalidad de ítems del TRC-13.

Diferencias en Función del Sexo en el Tiempo de Respuesta Invertido en el TRC según la Condición Experimental

La Tabla 34 recoge las diferencias entre hombres y mujeres en el TR invertido en responder a cada ítem del test y al sumatorio de ítems que componen el TRC-13, TRC-10 y TRC-3, en cada una de las condiciones experimentales.

Como se puede observar, los datos muestran que, en cómputo general, no existen diferencias significativas entre hombres y mujeres en cuanto al tiempo que invierten en responder a los ítems del TRC. Únicamente se han encontrado que los hombres invierten más tiempo que las mujeres en la deliberación de los ítems originales (TRC-3) cuando estos se disponen al inicio del test (condición 1). Sin embargo, estas diferencias fueron pequeñas y marginalmente significativas ($d = .28$). En las otras dos condiciones, hombres y mujeres no difieren significativamente en cuanto al tiempo de reacción. Estos datos sugieren que, independientemente del orden de presentación de los ítems del TRC-13, hombres y mujeres invierten en promedio el mismo tiempo para deliberar y responder a los ítems.

Tabla 34.

Diferencias entre Hombres y Mujeres en el Tiempo de Respuesta Invertido en el TRC según la Condición Experimental

	Condición 1						Condición 2						Condición 3								
	Hombres			Mujeres			Hombres			Mujeres			Hombres			Mujeres					
	\bar{X}	SD		\bar{X}	SD	d	F			\bar{X}	SD	d	F			\bar{X}	SD	d	F		
Ítem 1	31.97	25.28		31.49	26.98	.02	.02			27.69	17.69	29.16	21.77	-.07	.23	33.08	29.85	32.67	25.42	.02	.01
Ítem 2	45.03	38.40		36.63	24.44	.28	3.50 ¹			32.78	29.61	31.67	25.81	.04	.08	30.84	26.09	29.17	26.00	.06	.21
Ítem 3	45.33	30.11		37.64	22.99	.30	4.01*			40.23	19.64	40.83	20.39	-.03	.04	41.63	29.62	39.19	26.76	.09	.38
Ítem 4	32.85	32.38		31.24	28.13	.05	.13			51.20	63.49	46.87	45.87	.08	.31	44.77	32.70	46.87	39.75	-.06	.16
Ítem 5	54.17	34.97		49.26	41.82	.12	.69			49.62	30.07	50.16	36.95	-.02	.01	61.52	44.60	52.89	47.15	.19	1.76
Ítem 6	56.58	46.00		49.81	38.39	.16	1.21			64.95	52.66	66.55	67.59	-.03	.03	69.78	49.48	60.80	45.04	.19	1.84
Ítem 7	41.67	29.90		41.73	32.58	.00	.00			40.18	20.58	39.72	23.80	.02	.02	51.27	33.35	42.92	29.84	.27	3.55
Ítem 8	67.45	63.39		57.54	46.66	.19	1.56			68.82	48.70	55.04	43.82	.30	4.10*	72.53	46.57	58.52	43.43	.31	4.90
Ítem 9	36.38	21.01		34.83	22.18	.07	.23			39.64	22.39	37.01	19.66	.13	.73	46.39	30.27	40.23	25.87	.22	2.47
Ítem 10	72.95	54.71		65.29	48.20	.15	1.03			62.12	40.70	66.42	47.61	-.10	.40	74.26	64.24	70.19	49.79	.07	.26
Ítem 11	57.85	46.76		69.70	51.43	-.24	2.54			55.45	31.16	68.00	49.27	-.29	3.65 ¹	64.29	43.13	69.28	46.62	-.11	.61
Ítem 12	35.72	27.28		34.38	29.16	.05	.10			48.28	37.13	33.78	25.81	.48	10.36**	46.35	36.95	41.44	32.26	.14	1.03
Ítem 13	89.50	74.50		94.91	79.30	-.07	.22			93.25	73.35	90.52	82.07	.03	.05	87.63	68.31	87.72	65.77	.00	.00
TRC-13	667.44	284.88		634.46	306.70	.11	.54			674.23	249.15	655.74	284.78	.07	.21	724.35	321.43	671.90	272.96	.18	1.59
TRC-3	122.32	71.41		105.76	51.02	.28	3.54 ¹			100.70	44.10	101.66	46.67	-.02	.02	105.55	62.52	101.03	55.29	.08	.300
TRC-10	545.11	237.28		528.71	275.10	.06	.18			573.53	220.43	554.08	256.09	.08	.28	618.79	273.85	570.86	238.83	.19	1.78

Nota. Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; SD = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = sumatorio de los ítems originales (Frederick, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el 10.

¹ $p \leq .10$; * $p \leq .05$; ** $p \leq .01$

Relación entre el Desempeño en el TRC y el Tiempo Invertido en Responder

El tercer objetivo de este estudio era examinar la relación entre el rendimiento en el TRC y el tiempo invertido en responder al test. Para ello se calcularon las correlaciones policóricas observadas entre las variables y las correlaciones corregidas por atenuación en la fiabilidad de ambas variables. Los coeficientes de fiabilidad por consistencia interna del tiempo invertido en responder al TRC-13, a los nuevos ítems (TRC-10) y a los ítems originales (TRC-3) fueron calculados como alfas de Cronbach. Las magnitudes de esos coeficientes para el total de la muestra y para la muestra de hombres y mujeres están recogidos en la Tabla 35.

Tabla 35.

Coefficientes de Fiabilidad del Tiempo de Respuesta en el TRC-13, TRC-10 y TRC-3

	TR-13	TR-10	TR-3
Muestra total	.785	.750	.492
Hombres	.778	.729	.539
Mujeres	.789	.762	.450

Nota. TR-13 = tiempo de respuesta en el TRC-13; TR-10 = tiempo de respuesta en el TRC-10; TR-3 = tiempo de respuesta en el TRC-3.

Los coeficientes de fiabilidad del TRC-13 y del TRC-10 utilizados fueron los coeficientes de fiabilidad *CES* calculados en el estudio 1 (a saber, $CES_{TRC-13} = .66, .67$ y $.61$ para la muestra total, hombres y mujeres, respectivamente y $CES_{TRC-10} = .57, .52$ y $.53$, respectivamente para la muestra total, hombres y mujeres). Asimismo, el coeficiente de fiabilidad del TRC-3 fue el coeficiente test-retest también calculado en el estudio 1 ($r_{xx} = .71, .76, .69$ para la muestra total, hombres y mujeres, respectivamente). La Tabla 36 recoge las correlaciones obtenidas entre las variables.

Como se puede observar, la relación entre el rendimiento en el TRC y el tiempo invertido en responder al test para el total de la muestra fue positiva y significativa en todos los casos. Los tamaños del efecto corregidos encontrados fueron .328, .315 y .240, respectivamente para el TRC-13, TRC.10 y TRC-3. Por lo tanto, estos datos confirman la Hipótesis 2 de este estudio.

Respecto a los resultados en función del sexo, los datos muestran una relación positiva y significativa en el caso de las mujeres (rango de $\rho = .375$ a $.515$) y prácticamente nula en el caso de los hombres (rango de $\rho = -.125$ a $.061$). Así pues, los datos sugieren que, en el caso de las mujeres, invertir más tiempo en responder a los ítems podría suponer una mejora en los resultados en el test, mientras que, en el caso de los hombres, invertir más tiempo en la deliberación de las respuestas no tiene porqué producir efectos en su rendimiento.

Tabla 36.

Correlaciones Policóricas entre los TRC y el TR

	TR-13	TR-3	TR-10
<i>Muestra Total</i>			
TRC-13	.236** (.328**)	.101* (.177**)	.236** (.333**)
TRC-3	.257** (.344**)	.142** (.240**)	.259** (.368**)
TRC-10	.257** (.384**)	.084 (.159**)	.207** (.315**)
TR-3	.697**	-	-
TR-10	.987**	.577**	-
<i>Hombres</i>			
TRC-13	.006 (.008)	-.050 (-.083)	.012 (.017)
TRC-3	.029 (.038)	-.047 (-.073)	.045 (.061)
TRC-10	-.010 (-.016)	-.066 (-.125)	-.006 (-.010)
TR-3	.691**	-	-
TR-10	.986**	.561**	-
<i>Mujeres</i>			
TRC-13	.357** (.515**)	.212** (.405**)	.345** (.506**)
TRC-3	.362** (.491**)	.249** (.447**)	.351** (.478**)
TRC-10	.317** (.490**)	.183** (.375**)	.305** (.480**)
TR-3	.699**	-	-
TR-10	.988**	.585**	-

Nota. Las correlaciones entre paréntesis son correlaciones corregidas por atenuación en la fiabilidad de ambas variables. TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = TRC original (Frederick, 2005); TRC-10 = ítems desarrollados por Salgado (2014a) eliminando el ítem 10; TR-13 = tiempo de respuesta del TRC-13; TR-3 = tiempo de respuesta del TRC-3; TR-10 = tiempo de respuesta del TRC-10.

* $p \leq .05$; ** $p \leq .001$

4. DISCUSIÓN

Los objetivos de este estudio han sido 4. En primer lugar, se ha querido explorar el efecto que el orden de presentación de los ítems tiene sobre los resultados en un nuevo TRC (Salgado, 2014a).

En relación con esto, los resultados han mostrado que cuando los ítems originales se encuentran posicionados en medio de la secuencia de los nuevos ítems (condición 2), los participantes alcanzan significativamente mejores resultados en los ítems originales (TRC-3) y en el sumatorio total del TRC-13 que cuando se encuentran posicionados al inicio (condición 1) o al final (condición 3) de la secuencia de ítems. Del mismo modo, cuando los ítems originales se disponen al final de la secuencia de ítems (condición 3), el rendimiento en estos ítems también se ve mejorado significativamente, respecto a posicionarlos al inicio de la secuencia (condición 1); aunque, en este caso, no se produce un efecto sobre el rendimiento en el total del TRC-13. En ambos casos, el rendimiento en los nuevos ítems (TRC-10) no se ve significativamente afectado por el orden de presentación de los ítems.

Por lo tanto, la primera conclusión extraída de este estudio es que el orden de presentación de los ítems del TRC-13 tiene un efecto sobre el rendimiento en el test. Disponiendo los 3 ítems originales de Frederick (2005; Kahneman y Frederick, 2002) en medio de la secuencia de ítems desarrollados por Salgado (2014a) se puede maximizar el rendimiento de los participantes en el total del test y en los ítems originales. De modo que el orden de presentación de los ítems parece no producir un impacto en el rendimiento alcanzado en los nuevos ítems.

El segundo objetivo del estudio ha sido investigar si el orden de presentación de los ítems tiene un efecto sobre el tiempo de respuesta al test. En este sentido, los resultados han indicado que el orden en que los ítems son dispuestos en el TRC-13 no produce efectos en el tiempo que los participantes invierten en responder al conjunto de ítems del TRC-13. Sin embargo, los resultados sugieren que los participantes tienden a invertir significativamente más tiempo en la deliberación de los ítems originales o de los nuevos ítems cuando estos se encuentran posicionados al inicio de la secuencia de ítems (condición 1 para los ítems originales y condición 3 para los nuevos ítems). Este hallazgo sugiere que en el proceso de respuesta al test se puede estar dando un *efecto priming*. Es decir, cuando los participantes comienzan a responder al TRC, invierten más tiempo en la deliberación de los primeros ítems por la dificultad que estos entrañan y por la falta de familiaridad con la mecánica de respuesta. Sin embargo, la resolución de los primeros ítems da paso a la experiencia y a la automaticidad del proceso de respuesta, por lo que el tiempo invertido en responder a los siguientes ítems se reduce significativamente.

No obstante, los resultados sobre los efectos que el orden de presentación de los ítems tiene en el rendimiento en el test han indicado que, a pesar de que los participantes invierten más tiempo en la deliberación de los ítems originales, cuando estos se sitúan al inicio del TRC-13, también alcanzan significativamente un peor desempeño en esos ítems (pero no en los nuevos

ítems). Estos resultados podrían sugerir que el éxito en el desempeño de los ítems originales podría venir determinado por una mayor dificultad. Los resultados de este estudio han mostrado que una proporción mayor de participantes fallan al responder a los ítems de Frederick en comparación con responder a los ítems de Salgado (2014a). Además, como previos autores han señalado, el éxito en el desempeño de esos ítems también podría estar determinado por un problema de *mindware*, es decir, por la falta de conocimiento declarativo o de reglas estratégicas necesarias para resolver los problemas (Sinayev y Peters, 2015; Szaszi et al., 2017). Por ello, situando estos ítems en medio o al final del TRC-13, se puede conseguir una mejora del rendimiento a través del *efecto priming* que implica practicar previamente con otros problemas más fáciles que, además, envuelven la misma mecánica de respuesta que los originales. En cualquier caso, nuevos estudios deberían seguir examinando estas cuestiones con mayor detenimiento.

El tercer objetivo de este estudio ha sido determinar la relación entre el tiempo de respuesta y el desempeño en el TRC. El análisis de correlaciones bivariadas ha mostrado una relación positiva entre el desempeño en el TRC y el tiempo de respuesta invertido.

Finalmente, el cuarto objetivo ha sido explorar los efectos del orden de presentación de los ítems del TRC sobre los resultados en el test y sobre el tiempo de respuesta en hombres y mujeres de forma independiente. En este sentido, los datos muestran resultados muy similares a los encontrados en la muestra total.

En relación con el desempeño en el test, cuando se sitúan los ítems originales en el medio de los nuevos ítems (condición 2), tanto los hombres como las mujeres alcanzan significativamente mejores resultados en los ítems originales (TRC-3) que cuando se sitúan en cualquiera de las otras dos combinaciones (al inicio o al final del test) y, únicamente, en el caso de las mujeres también se incrementa significativamente el desempeño total en el TRC-13. Del mismo modo, el rendimiento de las mujeres en los nuevos ítems (TRC-10) también se incrementa significativamente cuando los ítems originales se sitúan en medio de la secuencia que cuando se disponen al final de la misma.

En relación con el tiempo de respuesta, el estudio muestra que el orden de presentación de los ítems del TRC no tiene efectos sobre el tiempo que las mujeres invierten en responder al TRC-13 y a los ítems originales y nuevos por separado. Pero, sin embargo, si existe una correlación positiva entre el tiempo de respuesta y el resultado en el test. En el caso de los hombres, los resultados replican los obtenidos en la muestra total, de modo que, únicamente, invierten más tiempo en la deliberación de los ítems (originales o nuevos) cuando estos se sitúan al inicio del TRC (condición 1 y 3, respectivamente).

Finalmente, este estudio no ha encontrado diferencias significativas entre hombres y mujeres en cuanto al tiempo invertido en responder al TRC-13, a los ítems originales o a los nuevos ítems, en ninguna de las condiciones experimentales. Sin embargo, se ha encontrado que,

en todos los casos, los hombres alcanzan sistemáticamente mejores resultados en el test que las mujeres. Esto sugiere que las diferencias en función del sexo en el desempeño en el test podrían no estar determinadas por el tiempo invertido en responder a los ítems. Igualmente, también se ha encontrado que las diferencias en función del sexo en la RC fueron mayores en la condición 2 que en las otras dos condiciones y justamente ambos grupos alcanzaron mejores resultados en el test en esa condición. Esto podría sugerir que maximizar el desempeño de hombres y mujeres en el TRC no implicaría un incremento de las diferencias en el desempeño en el test entre ambos grupos.

Al igual que cualquier estudio experimental, esta investigación también presenta algunas limitaciones. En primer lugar, los participantes han sido incentivados con el objetivo de motivarlos a participar en el estudio y a maximizar su desempeño en la tarea. No obstante, el tipo de incentivo ofrecido ha variado en algunos casos. Así pues, se han ofrecido incentivos económicos y no económicos (informe de los resultados) e incluso dentro de los incentivos económicos se han compensado con diferentes cantidades monetarias, lo que podría afectar a los resultados alcanzados. Por otro lado, el tamaño de la muestra de hombres en cada una de las condiciones experimentales es reducido, especialmente si lo comparamos con el tamaño de la muestra de las mujeres en esas mismas condiciones experimentales. Esto implica que los resultados relacionados con los hombres pueden estar afectados por mayor error de muestreo. Finalmente, como se ha mostrado en el estudio 1, el desempeño de los sujetos en el TRC puede estar determinado por el formato de administración del test. De este modo, los resultados encontrados deberían ser replicados en otras investigaciones en las que el TRC sea administrado en formato papel y lápiz.

Para resumir, este estudio nos ha permitido examinar los efectos que tiene la manipulación del orden de presentación de los ítems en el desempeño en el test y en el tiempo de respuesta al mismo. Asimismo, hemos podido examinar la relación entre el tiempo de reacción y el desempeño no solo para el total de la muestra sino también para los grupos de hombres y de mujeres de forma independiente. Los resultados encontrados han reflejado que presentar los ítems originales en medio de los nuevos mejora significativamente el desempeño en el TRC-13 y en los ítems originales. Asimismo, se ha encontrado que el orden de presentación de los ítems no afecta al tiempo invertido en responder al TRC-13. Aun así, los datos parecen indicar que los participantes tienden a invertir significativamente más tiempo en la deliberación de los ítems (originales o nuevos) cuando estos se encuentran al inicio del test.



ESTUDIO 3.

Validez de Constructo y de Criterio del Test de Reflexibilidad Cognitiva

1. INTRODUCCIÓN

La reflexibilidad cognitiva (RC) es un concepto reciente en la literatura científica. Se han llevado a cabo algunas investigaciones con el objetivo de examinar su validez de constructo, sin embargo, muy pocos esfuerzos se han centrado en analizar su validez de criterio. A pesar de ello, sigue habiendo falta de consenso sobre qué mide realmente el TRC. Por ello, resulta necesario continuar estudiando y mostrando evidencias de validez de constructo y de criterio de la medida de RC. El principal objetivo de este estudio ha sido explorar ambas fuentes de validez.

En relación con la validez de constructo, se pueden distinguir tres principales propuestas sobre el propósito de este instrumento. La primera sostiene que el TRC es una medida de capacidad mental general y más concretamente de capacidad numérica (Sinayev y Peters, 2015; Weller et al., 2013). La segunda indica que el TRC, además de CMG, es un test que mide la capacidad o disposición hacia un pensamiento racional, entendido como la capacidad o disposición a inhibir el procesamiento miserable de nuestro sistema 1 y la tendencia a activar el procesamiento deliberado del sistema 2 (Frederick, 2005; Toplak et al., 2011, 2014). En una interpretación amplia de esta segunda hipótesis, se ha formulado una tercera propuesta: algunos autores defienden que el TRC, no solo mide capacidad cognitiva, sino también la disposición hacia un pensamiento racional de mente abierta, es decir, la disposición hacia un pensamiento minucioso, reflexivo y de cognición elaborada. Por lo tanto, identifican a la RC con la noción de *reflectiveness* o de *actively open-minded thinking* (AOT) de Baron (1988). En relación con esto, Juanchich y colaboradores (2016) añaden que el TRC, además de evaluar la tendencia a ser exhaustivo y/o a participar activamente en el pensamiento de mente abierta, mide un amplio conjunto de disposiciones. Los autores afirman que el TRC mide la capacidad más genérica de

usar la intuición cuando ésta sea apropiada, es decir, identifican la RC como un elemento que permite seleccionar el proceso cognitivo más adecuado a la situación específica.

En relación con la validez de criterio, muy pocos estudios han explorado esta cuestión. Así, mientras la validez para predecir criterios académicos ha sido ligeramente más investigada (Insler et al., 2015; Thomson y Oppenheimer, 2016), únicamente un estudio previo ha examinado la validez para predecir criterios organizacionales (Salgado et al., 2019).

En base a lo expuesto, el propósito de este estudio ha sido doble. Por un lado, se han tratado de mostrar evidencias de validez de constructo investigando la validez convergente y discriminante del TRC. Para ello: (1) se ha estudiado su relación con la capacidad mental general (CMG), con los cinco grandes factores de personalidad y con varias medidas de disposición de pensamiento. Asimismo, con el objetivo de comprobar cuáles de esas variables individuales, en conjunto, determinan la RC, (2) se ha explorado la predicción múltiple de la RC. Además, estudios previos han sugerido que los hombres alcanzan significativamente mejores resultados en el TRC que las mujeres (Brañas-Garza et al., 2015; Cueva et al., 2015; Frederick, 2005; Primi et al., 2018). Por lo que las diferencias en función del sexo en la RC podrían estar explicadas por la diferente capacidad de las variables disposicionales para determinar la RC en ambos grupos. El estudio de Campitelli y Gerrans (2014), mostró que las variables individuales que explicaron el desempeño en el TRC en hombres y mujeres no fueron exactamente las mismas y que, en los casos en que coincidieron, contribuyeron de forma diferente. En base a ello, (3) este estudio también examina la validez de constructo de la medida de RC de forma independiente en la muestra de hombres y de mujeres.

El segundo propósito de este estudio ha sido mostrar evidencias de validez de criterio, investigando la capacidad de la RC para predecir diversos criterios académicos y ocupacionales. Finalmente, se ha examinado si la RC añade validez sobre otras variables predictoras para explicar el éxito académico y ocupacional.

Capacidad Mental General (CMG)

Varios estudios han sugerido que existe una relación moderada del TRC con la capacidad mental general y con varias medidas de capacidades cognitivas específicas (Frederick, 2005; Koscielniak et al., 2016; Liberali et al., 2012; Teovanović et al., 2015; Trippas et al., 2015; Welsh et al., 2013), indicándose que el TRC mantiene una relación con inteligencia fluida y con inteligencia cristalizada (Corgnet, Espín y Hernán-González, 2015; Finucane y Gullion, 2010). En base a ello, presentamos la siguiente hipótesis:

Hipótesis 1: *La RC se relaciona moderada y positivamente con la CMG.*

Personalidad: Modelo de los Big Five Factors

La relación entre el TRC y los cinco factores de personalidad no ha sido muy investigada. A pesar de ello, los resultados parecen indicar que el constructo de RC es independiente de la personalidad. Uno de los factores más examinado ha sido la dimensión de apertura a la experiencia (AP). Este factor se define por la imaginación activa, la sensibilidad estética, la atención a los sentimientos internos y la preferencia por la variedad, curiosidad intelectual e independencia de juicio (Barrick y Mount, 1991; Digman, 1990; Goldberg, 1992; Hogan, 1991; Salgado, 1998). Es el factor de personalidad que más se asocia con CMG (Ackerman y Heggstad, 1997; Ashton et al., 2000; Bates y Shieles, 2003), lo que parece sugerir una posible relación con la RC. En este sentido, los resultados previos han mostrado correlaciones muy variadas entre la RC y el factor de AP. Así, algunos estudios han reflejado tamaños del efecto pequeños, pero significativos (Cokely et al., 2012; Corgnet, Hernán-Gonzalez y Mateo, 2015; Cueva et al., 2015; Teovanović et al., 2015), mientras que otros han reflejado correlaciones nulas o prácticamente nulas entre ambos constructos (Browne et al., 2014; Juanchich et al., 2016; Poore et al., 2014; Shenhav et al., 2011).

El factor de extraversión (EX) se refiere al grado de socialización de las personas, es decir, si les gusta la excitación, la muchedumbre, la actividad social; si son sociables, habladores y activos o por el contrario son más introvertidos, reservados e independientes (Barrick y Mount, 1991; Digman, 1990; Goldberg, 1992; Hogan, 1991; Salgado, 1998). Las personas que son extrovertidas necesitan constantemente de estímulos externos y por lo tanto les puede resultar más difícil mantener su atención prolongada sobre una determinada tarea. En base a ello, los sujetos extrovertidos podrían distraerse más fácilmente de las tareas que requieran RC y cometer más errores intuitivos. Resultados previos sobre la relación entre este factor y la RC han mostrado evidencias de una correlación negativa y muy pequeña o incluso nula (Cokely et al., 2012; Corgnet, Hernán-Gonzalez y Mateo, 2015; Cueva et al., 2015; Juanchich et al., 2016; Poore et al., 2014; Shenhav et al., 2011; Yilmaz y Saribay, 2016).

Por su parte, el factor de estabilidad emocional (EE) hace referencia al ajuste emocional o neuroticismo que presenta una persona, es decir, la tendencia general a experimentar afectos negativos tales como ira, disgusto, enfado o culpa (Barrick y Mount, 1991; Salgado, 1998). Una persona baja en EE puede no ser capaz de controlar sus emociones ante tareas que puedan suscitarle algún tipo de afecto negativo (como, por ejemplo, la ansiedad por los cálculos matemáticos que implica el TRC), provocando un peor desempeño en la tarea. No obstante, las evidencias empíricas encontradas hasta el momento muestran una nula asociación entre este factor y la RC (Agnew y Harrison, 2017; Corgnet, Hernán-Gonzalez y Mateo, 2015; Cueva et al., 2015; Juanchich et al., 2016; Poore et al., 2014; Yilmaz y Saribay, 2016).

El factor de amigabilidad (AM) hace referencia a las relaciones interpersonales del sujeto, de modo que, las personas amigables son altruistas, muestran simpatía por los demás y tratan de ayudarlos; mientras que, las personas antagonistas son egocéntricas, muestran una mayor predisposición a desconfiar de las intenciones de los demás y suelen ser competitivas (Barrick y Mount, 1991; Salgado, 1998). Los resultados previos sobre este factor indican una asociación nula o prácticamente nula, y en este caso, negativa con la RC (Corgnet, Hernán-Gonzalez y Mateo, 2015; Cueva et al., 2015; Juanchich et al., 2016; Poore et al., 2014; Yilmaz y Saribay, 2016).

Finalmente, el factor conciencia (C) es la dimensión relacionada con el control de los impulsos, es decir, con la organización y planificación de las tareas para alcanzar logros (Barrick y Mount, 1991; Salgado, 1998). En base a esto, parece lógico pensar que las personas más concienzudas y, por ende, con mayor control de sus impulsos, serán más capaces de resistirse a responder con la respuesta inmediata, dando paso a una mayor deliberación en el desempeño en el TRC. Sin embargo, los resultados encontrados hasta el momento parecen indicar una relación pequeña entre el factor conciencia y la RC, encontrándose en algunos casos, correlaciones negativas entre ambas variables (Corgnet, Hernán-Gonzalez y Mateo, 2015; Cueva et al., 2015; Juanchich et al., 2016; Poore et al., 2014; Yilmaz y Saribay, 2016).

En base a las evidencias expuestas, las hipótesis planteadas fueron:

Hipótesis 2a: *La RC se relaciona positivamente con el factor de AP.*

Hipótesis 2b: *La RC no correlaciona con los factores de personalidad de EE, EX, AM y C.*

Disposiciones de Pensamiento (DP)

Diversos estudios han examinado la relación entre el TRC y las medidas de estilos cognitivos. La disposición hacia un pensamiento racional ha sido una de las más ampliamente exploradas. En este sentido, se ha encontrado una asociación negativa entre la RC y la tendencia hacia un pensamiento avaro, medido éste mediante la comisión de sesgos cognitivos (Aczel et al., 2015; Albaity et al., 2014; Cesarini et al., 2012; Frederick, 2005; Hardisty y Weber, 2009; Hoppe y Kusterer, 2011; Oechssler et al., 2009; Toplak et al., 2011, 2017).

Una segunda disposición de pensamiento ampliamente asociada con el TRC ha sido la disposición hacia un pensamiento minucioso, cuidadoso y de cognición elaborada. Este estilo cognitivo se asocia con el amplio concepto de *Actively Open-Minded Thinking* (AOT; Baron, 1988). Se ha encontrado una relación entre el TRC y AOT (Haran et al., 2013; Svedholm-Häkkinen y Lindeman, 2018; Thompson et al., 2013; Toplak et al., 2014, 2017).

Asimismo, se ha explorado la relación entre la RC y otras disposiciones de pensamiento. Los resultados encontrados han sugerido la existencia de una relación positiva con la disposición hacia un pensamiento de compromiso y disfrute con el esfuerzo cognitivo (Frederick, 2005; Haran et al., 2013; Thomson y Oppenheimer, 2016; Toplak et al., 2014) y con los 2 estilos de decisión intuitivo y espontáneo (Juanchich et al., 2016). Asimismo, otros estudios hallaron una relación con el pensamiento orientado al futuro (Toplak et al., 2014, 2017), con la evitación de un pensamiento supersticioso (Toplak et al., 2014, 2017) y con el pensamiento hacia la creencia en soluciones simples (van Prooijen, 2017).

En base a estas evidencias, se propone la siguiente hipótesis:

Hipótesis 3: *La RC presenta una relación con los estilos cognitivos analíticos.*

Desempeño

La validez de criterio del TRC no ha sido investigada en profundidad. Existen relativamente pocos estudios que exploran esta cuestión en el ámbito académico y, a nuestro juicio, solo existe un trabajo publicado en relación con el desempeño ocupacional. Por lo tanto, este estudio también trata de examinar la validez predictiva del TRC y mostrar nuevos indicios al respecto.

La investigación previa ha reflejado correlaciones positivas entre la RC y el desempeño académico. Concretamente con medidas tales como GPA (*Grade Point Average*; Agnew y Harrison, 2017; Borghans y Golsteyn, 2014; Insler et al., 2015; Thomson y Oppenheimer, 2016; Toplak et al., 2014), logro matemático (Gómez-Chacón et al., 2014; Lindeman y Svedholm-Häkkinen, 2016) y de las calificaciones de exámenes, como por ejemplo, de física (Lindeman y Svedholm-Häkkinen, 2016), de nivel A de alemán (Lohse, 2016) y de estadística (Primi et al., 2017). Asimismo, se han encontrado evidencias de que la RC no añade validez sobre otras variables individuales para predecir el desempeño académico (Gómez-Chacón et al., 2011; Toplak et al., 2014). En conjunto, estas evidencias sugieren que la RC puede ser un predictor válido de diversas medidas de desempeño académico.

Respecto al desempeño ocupacional, existe muy poca evidencia de la validez de la RC para predecir el desempeño ocupacional, aun así, los indicios sugieren que la RC podría ser un predictor válido de criterios ocupacionales (Salgado et al., 2019).

En base a las evidencias empíricas expuestas se proponen las hipótesis:

Hipótesis 4a: *La RC es un predictor válido del desempeño académico.*

Hipótesis 4b: *La RC es un predictor válido del desempeño ocupacional.*

Finalmente, este estudio también ha examinado si la RC añade validez sobre otras variables individuales para predecir las variables criterio examinadas. Los estudios previos han encontrado que el TRC no añade validez sobre otras medidas para pronosticar el desempeño académico (Gómez-Chacón et al., 2011; Toplak et al., 2014), pero, sin embargo, sí se ha encontrado que contribuye en la explicación del desempeño ocupacional (Salgado et al., 2019). En este sentido, nuestro estudio será meramente exploratorio y no proponemos hipótesis al respecto.

2. MÉTODO

Muestra

La muestra empleada en este estudio es la misma utilizada en el estudio 1. Está compuesta por 1,367 sujetos, de los cuales el 64.16% eran mujeres ($n = 877$) y la edad promedio fue de 21.72 años ($SD = 4.14$, rango = 17 a 58 años). Todos los participantes eran estudiantes de la Universidad de Santiago de Compostela (USC) y el 15.56% de ellos eran postgraduados. El restante 84.44% estaba formado por alumnos matriculados en diversas titulaciones de la USC. La Tabla 37 recoge las frecuencias de los estudiantes según el área de estudio. La muestra total se compuso de pequeñas submuestras reclutadas, a través de diferentes métodos, durante los años 2015 a 2018. Así, el 31% de los sujetos ($n = 423$) fueron reclutados en las aulas y el restante 69% ($n = 944$) fueron participantes voluntarios de estudios experimentales.

Procedimiento

Para participar en los estudios experimentales, se impuso la condición de que los estudiantes debían entregar una copia de su expediente académico, donde se recogiesen las asignaturas superadas hasta el momento, la calificación obtenida en cada una de ellas y la nota media del expediente en base a 10. Asimismo, a un grupo más reducido de la muestra (14.19%) se les pidió que entregasen también una copia de la tarjeta de las calificaciones de la EBAU (Evaluación de Bachillerato para el Acceso a la Universidad, antiguamente conocida como PAU o Prueba de Acceso a la Universidad) en el que se recogiese la nota media de acceso a la universidad.

Con independencia del método de reclutamiento de las muestras, se incentivó la participación de los sujetos con el objetivo de maximizar su desempeño. El 55.74% de la muestra recibieron una compensación económica (entre 10 y 15€) a cambio de su participación, mientras que los restantes recibieron un informe con los resultados obtenidos en las diversas pruebas administradas. Asimismo, se instruyó a los participantes en el carácter confidencial y anónimo del estudio y en el único propósito e interés en examinar los resultados de modo agregado.

Tabla 37.

Características de la Muestra (N = 1,367)

	<i>n</i>	%
Sexo		
Mujeres	877	64.16
Hombres	490	35.84
Edad		
≤ 20	660	48.28
21-23	457	33.43
> 23	247	18.07
No indicado	3	.22
Área de Estudio		
Ciencias sociales	788	57.64
Ciencias experimentales	170	12.44
Ciencias de la salud	311	22.75
Humanidades	59	4.32
No indicado	39	4.85
Lugar de realización del estudio		
Aula	423	31
Laboratorio	944	69

Con el objetivo de calcular el coeficiente de fiabilidad test-retest de algunas de las medidas utilizadas se animó a los participantes de los estudios experimentales ($n = 944$) a participar de nuevo en un momento temporal posterior. Entre la realización del TRC en el momento temporal 1 y 2 transcurrieron, en promedio, 96.67 días ($SD = 170.90$). En este segundo caso, la muestra se redujo a 295 sujetos, de los cuales 210 fueron mujeres. La edad promedio de la muestra fue de 21.89 ($SD = 4.12$) y el rango de edad abarcó desde los 17 a los 58 años.

Instrumentos de Medida

Reflexibilidad Cognitiva (RC)

Test de Reflexibilidad Cognitiva (TRC-13; Salgado, 2014a; véase Apéndice D). Para responder a los 13 ítems que componen la medida, los participantes tuvieron que seleccionar una de las dos opciones de respuesta propuestas. No dispusieron de límite temporal para hacerlo y, por cada ítem resuelto correctamente, recibieron un punto.

En el apartado de resultados se examinan las evidencias de validez de la RC empleando el TRC-13, TRC-3 y TRC-10. Los coeficientes de fiabilidad de las tres medidas de RC para esta muestra son los descritos en el estudio 1 (véase Tabla 28). Los coeficientes de restricción en el rango del TRC para cada una de las medidas criterio de este estudio se presentan en la Tabla 38.

Tabla 38.

Coefficientes de Restricción en el Rango del TRC en Función de las Variables Criterio y las Muestras de Estudio

	TRC-13		TRC-10		TRC-3	
	<i>U</i>	<i>u</i>	<i>U</i>	<i>u</i>	<i>U</i>	<i>u</i>
<i>Muestra Total</i>						
CDTE	1.006	.994	1.010	.990	.994	1.006
CDCE	1.005	.995	1.010	.990	.991	1.009
CDAN	1.022	.978	1.032	.969	.997	1.003
GCT	1.064	.940	1.071	.934	1.025	.976
GPA	1.033	.968	1.041	.961	1.001	.999
Bachillerato	1.031	.970	1.038	.964	.991	1.009
CAU	1.042	.959	1.054	.948	.997	1.003
<i>Hombres</i>						
CDTE	1.043	.959	1.047	.955	1.023	.978
CDCE	1.040	.961	1.043	.959	1.020	.981
CDAN	1.072	.933	1.074	.931	1.049	.954
GCT	1.014	.986	1.019	.981	1.043	.958
GPA	1.080	.926	1.087	.920	1.038	.963
Bachillerato	.901	1.109	.888	1.126	1.005	.995
CAU	.955	1.047	.943	1.061	1.022	.979
<i>Mujeres</i>						
CDTE	1.005	.995	1.009	.991	.994	1.006
CDCE	1.003	.997	1.008	.992	.994	1.006
CDAN	1.022	.978	1.033	.968	.993	1.007
GCT	1.081	.925	1.089	.918	1.017	.983
GPA	1.029	.971	1.035	.966	1.000	1.000
Bachillerato	1.069	.935	1.087	.920	.973	1.028
CAU	1.062	.942	1.080	.926	.988	1.012

Nota. *U* = coeficiente de restricción en el rango; *u* = 1/*U*; CDTE = escala de desempeño de tarea; CDCE = escala de desempeño contextual; CDAN = escala de conductas contraproductivas; GCT = general clerical test; GPA = grade point average; CAU = calificaciones de acceso a la universidad.

Capacidad Mental General (CMG)

Para medir la capacidad mental general (en adelante CMG) de los sujetos se emplearon 3 medidas diferentes: (1) el Wonderlic Personnel Test (WPT) de Wonderlic (1998), (2) el Test 3 – Forma A del factor G de Cattell (Factor G; Cattell y Cattell, 1986) y (3) el test de Razonamiento Lógico Simbólico (RLS) de Metis (2001). A continuación, se sintetizan las principales características de estos tests.

Wonderlic Personnel Test (WPT; Wonderlic, 1998). El test mide inteligencia general o factor “g”, es decir, se emplea para describir la capacidad individual de aprendizaje, de entender instrucciones y resolver problemas. Se compone de 50 ítems y se define como un test ómnibus, donde los ítems incluyen problemas de diferentes tipos: series de números, análisis de figuras

geométricas, comparación de números, comparación de palabras, frases desordenadas y paralelismos en oraciones. Los ítems se disponen a lo largo del test en orden ascendente en dificultad. Al tratarse de una prueba de desempeño máximo, los participantes dispusieron de 12 minutos para completar el test. La puntuación final se calculó con el sumatorio de ítems respondidos correctamente. De modo que, las puntuaciones finales pueden oscilar entre 0 y 50. Puntuaciones altas en esta medida indican mayor CMG.

El 81.13% de los participantes del estudio cubrieron el WPT. La Tabla 39 muestra la fiabilidad de la medida para la muestra total y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad: dos por consistencia interna (*KR20* y la división en dos mitades, con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente *CES*. Asimismo, la Tabla 39 recoge el valor del coeficiente de restricción en el rango (*U* y *u*) calculado para la muestra del estudio.

Tabla 39.

Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango del WPT

	<i>KR20</i> (<i>n</i>)	División 2 mitades (<i>n</i>)	Test-retest (<i>n</i>)	<i>CES</i> (<i>n</i>)	<i>U</i>	<i>u</i>
Muestra total	.88 (1,109)	.91 (1,109)	.80 (288)	.85 (288)	1.36	.74
Hombres	.90 (374)	.92 (374)	.80 (86)	.84 (86)	1.35	.74
Mujeres	.87 (735)	.90 (737)	.78 (202)	.84 (202)	1.36	.74

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; *KR20* = coeficiente de fiabilidad por consistencia interna Kuder-Richardson 20; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; *CES* = coeficiente de estabilidad y equivalencia; *U* = coeficiente de restricción en el rango; *u* = $1/U$.

Test 3 – Forma A del factor G de Cattell (Factor G; Cattell y Cattell, 1986). Este test pertenece al tipo de pruebas “no verbales” o “test libres de cultura”, ya que para su realización solo se requiere que el sujeto perciba la diferencia entre formas y figuras. Esta prueba está formada por 4 subtests con contenidos perceptivos diferentes para evitar que determinadas diferencias en esta aptitud perceptiva influyan en el resultado de la medida final de CMG. El primer subtest consta de 13 series incompletas donde los participantes deben seleccionar la opción que continúe adecuadamente la serie. El segundo subtest es el de clasificación y consiste en seleccionar dos figuras que son discordantes respecto a las de su serie. El subtest matrices consiste en completar las diferentes matrices o cuadros de figuras escogiendo alguno de los dibujos propuestos como solución. Por último, en el subtest condiciones se debe escoger aquella alternativa que cumpla las mismas condiciones que la figura de referencia. En total, los participantes dispusieron de 12 minutos y medio para completar la prueba. La puntuación final del test fue la suma del número

de problemas resueltos correctamente. El rango de puntuaciones puede oscilar entre 0 a 50. Puntuaciones altas en esta medida también indican mayor CMG.

Del total de la muestra, únicamente se administró este test al 22.39%. La Tabla 40 muestra la fiabilidad de la medida para la muestra conjunta y para el grupo de hombres y mujeres. En este caso, se calcularon únicamente 2 coeficientes de fiabilidad al no disponer de resultados de esta medida en un momento temporal 2. Los coeficientes calculados fueron el coeficiente de consistencia interna Kuder- Richardson 20 y la división del test en dos mitades, con corrección por la fórmula de Spearman-Brown. Asimismo, la Tabla 40 recoge el valor del coeficiente de restricción en el rango (U y u) calculado para la muestra total del estudio. En este caso, no se pudieron calcular los coeficientes de restricción en el rango del grupo de hombres y mujeres por desconocer las desviaciones típicas de la población en ambos grupos.

Tabla 40.

Coefficientes de Fiabilidad y Coeficiente de Restricción en el Rango del Factor G de Cattell

	<i>KR20</i>	División 2 mitades	<i>n</i>	<i>U</i>	<i>u</i>
Muestra Total	.71	.79	330	.99	1.01
Hombres	.76	.82	139	-	-
Mujeres	.65	.76	191	-	-

Nota. *KR20* = coeficiente de fiabilidad por consistencia interna Kuder-Richardson 20; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; n = tamaño de la muestra; U = coeficiente de restricción en el rango; $u = 1/U$.

Test de Razonamiento Lógico Simbólico (RLS; Metis, 2001). Mide la capacidad individual para aplicar el razonamiento abstracto. Está compuesto por 25 ítems y se dispone de 15 minutos para completar el conjunto de problemas. La puntuación final del test se obtiene mediante el sumatorio del número de ítems resueltos correctamente. El rango de puntuaciones puede oscilar entre 0 y 25. Puntuaciones altas en esta medida indican mayor CMG.

Se administró este test al 22.39% de la muestra. La Tabla 41 recoge la fiabilidad de la medida para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 2 coeficientes de fiabilidad al no disponer de resultados de esta medida en un momento temporal 2. Los coeficientes calculados fueron el coeficiente de consistencia interna Kuder-Richardson 20 y la división del test en dos mitades con corrección por la fórmula de Spearman-Brown. Asimismo, la Tabla 41 recoge el valor del coeficiente de restricción en el rango (U y u) para las muestras de estudio.

Tabla 41.

Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango del RLS

	<i>KR20</i>	División 2 mitades	<i>n</i>	<i>U</i>	<i>u</i>
Muestra Total	.69	.65	306	1.03	.97
Hombres	.74	.71	118	1.19	.84
Mujeres	.59	.55	188	.90	1.10

Nota. *KR20* = coeficiente de fiabilidad por consistencia interna Kuder-Richardson 20; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; *n* = tamaño de la muestra; *U* = coeficiente de restricción en el rango; *u* = $1/U$.

Personalidad: Modelo de los Big Five Factors

Para evaluar la personalidad de los sujetos se emplearon 2 medidas diferentes: (1) el Inventario IP/5F de Salgado (1998) y (2) el Inventario QI5F-Tri de Salgado (2014c). Ambas medidas evalúan la personalidad normal de los individuos en base a los 5 grandes factores de personalidad: estabilidad emocional, extraversión, apertura a la experiencia, amigabilidad y conciencia (Costa y McCrae, 1988; Goldberg, 1992; Hogan, 1991).

Inventario de Personalidad IP/5F (IP/5F; Salgado, 1998). Es una medida de respuesta tradicional o *single stimulus* que está compuesta por 200 ítems. Cada uno de los cinco factores de personalidad son examinados por un conjunto de 40 ítems. Para responder a los ítems, los participantes deben indicar su grado de acuerdo con cada una de las afirmaciones presentadas. Para ello deben emplear una escala de 3 puntos: en *desacuerdo*, *ni en acuerdo ni en desacuerdo* y *de acuerdo*. Un ejemplo de ítem de cada uno de los factores sería: (a) “*nunca me pongo nervioso por nada*” (EE), (b) “*los que me conocen piensan que soy alegre y vitalista*” (EX), (c) “*soy una persona muy imaginativa*” (AP), (d) “*Creo que en general la gente es honesta y de buenas intenciones*” (AM) y (e) “*todo lo hago concienzudamente*” (C).

El inventario de personalidad IP/5F fue administrado al 47.40% de la muestra. La Tabla 42 refleja la fiabilidad de cada uno de los factores para la muestra conjunta y para el grupo de hombres y mujeres. En este caso, se calculó el coeficiente de fiabilidad por consistencia interna Alfa de Cronbach. Asimismo, la Tabla 42 recoge los valores del coeficiente de restricción en el rango (*U* y *u*) para cada uno de los factores en función de la muestra de estudio.

Tabla 42.

Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango de los Cinco Factores de Personalidad del Inventario IP/5F

IP/5F	α	n	U	u
<i>Muestra total</i>				
Estabilidad Emocional	.91	648	1.07	.94
Extroversión	.85	648	1.13	.89
Apertura Experiencia	.84	648	1.01	.99
Amigabilidad	.78	648	.91	1.11
Conciencia	.85	648	1.05	.95
<i>Hombres</i>				
Estabilidad emocional	.89	253	1.07	.94
Extroversión	.84	253	1.15	.87
Apertura Experiencia	.85	253	.98	1.02
Amigabilidad	.81	253	.84	1.19
Conciencia	.85	253	1.08	.92
<i>Mujeres</i>				
Estabilidad emocional	.90	395	1.07	.94
Extroversión	.86	395	1.12	.89
Apertura Experiencia	.84	395	1.04	.96
Amigabilidad	.76	395	.93	1.07
Conciencia	.85	395	1.02	.98

Nota. α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; n = tamaño de la muestra; U = coeficiente de restricción en el rango; $u = 1/U$.

QI5F-Tri (Salgado, 2014c). La segunda medida de personalidad empleada fue el inventario de personalidad quasi-ipsativo diseñado por Salgado (2014c). Se trata de un instrumento de respuesta de elección forzosa (*forced choice*) que mide los cinco factores de personalidad a través de 140 tríadas (28 tríadas por cada uno de los factores). Cada una de las tríadas engloba 3 rasgos de personalidad relacionados con diferentes factores de personalidad con el mismo grado de deseabilidad social. Los sujetos deben escoger, de entre los tres rasgos propuestos, el que mejor los define. A continuación, deben señalar entre los otros dos el que peor los define, de modo que al final, uno de los rasgos de la tríada quedará sin marcar. Un ejemplo de tríada del inventario de personalidad QI5F-Tri sería: “*soy una persona: (a) con grandes aptitudes, (b) me encantan los deportes acuáticos y (c) aprovecho todas las oportunidades para aprender*”.

El inventario de personalidad QI5F-Tri fue administrado al 64.89% de la muestra. La Tabla 43 recoge la fiabilidad de cada uno de los factores para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad: dos por consistencia interna

(Alfa de Cronbach y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente CES.

Tabla 43.

Coeficientes de Fiabilidad de los Cinco Factores de Personalidad de la Escala Q15F-Tri

	α (n)	División 2 mitades (n)	Test-retest (n)	CES (n)
<i>Muestra total</i>				
Estabilidad Emocional	.66 (887)	.56 (887)	.75 (128)	.55 (128)
Extroversión	.75 (887)	.75 (887)	.85 (128)	.67 (128)
Apertura Experiencia	.83 (887)	.84 (887)	.82 (128)	.71 (128)
Amigabilidad	.69 (887)	.69 (887)	.71 (128)	.58 (128)
Conciencia	.77 (887)	.72 (887)	.83 (128)	.61 (128)
<i>Hombres</i>				
Estabilidad emocional	.65 (261)	.54 (261)	.79 (32)	.66 (32)
Extroversión	.75 (261)	.75 (261)	.88 (32)	.77 (32)
Apertura Experiencia	.84 (261)	.86 (261)	.88 (32)	.82 (32)
Amigabilidad	.72 (261)	.75 (261)	.62 (32)	.52 (32)
Conciencia	.75 (261)	.71 (261)	.83 (32)	.62 (32)
<i>Mujeres</i>				
Estabilidad emocional	.64 (626)	.54 (626)	.71 (96)	.47 (96)
Extroversión	.76 (626)	.75 (626)	.84 (96)	.62 (96)
Apertura Experiencia	.83 (626)	.83 (626)	.81 (96)	.68 (96)
Amigabilidad	.68 (626)	.66 (626)	.75 (96)	.61 (96)
Conciencia	.77 (626)	.72 (626)	.83 (96)	.61 (96)

Nota. α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; n = tamaño de la muestra; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; CES = coeficiente de estabilidad y equivalencia; U = coeficiente de restricción en el rango; $u = 1/U$.

Disposiciones de Pensamiento (DP)

Para medir las disposiciones de pensamiento o estilos cognitivos de los sujetos se emplearon 4 medidas diferentes: (1) la Escala Breve de Autocontrol de Tangney et al. (2004), (2) la Escala *Need for Cognition* (NFC) de Salgado (2014b), (3) el Test de Figuras Enmascaradas (TFE) de Witkin, Oltman, Raskin y Karp (1971) y (4) la Escala de Conductas Personales (CP5) de Salgado (2014d). A continuación, se sintetizan las principales características de cada una de ellas.

Escala Breve de Autocontrol (Tangney et al., 2004). Mide la capacidad individual de anular o cambiar las respuestas internas, así como de interrumpir las tendencias de comportamiento no deseadas y frenar el impulso a actuar en base a ellas. Se trata, por tanto, de la capacidad individual de encontrar el mejor ajuste entre uno mismo y el entorno que le rodea (Tangney et al., 2004). Esta medida está compuesta por 13 ítems que tratan de medir diferentes

esferas del autocontrol: (1) el control de los pensamientos y las emociones, (2) el control de los impulsos, (3) la regulación del desempeño y (4) el control de malos hábitos. Cada uno de los ítems expresa una afirmación que debe ser respondida en función del grado de acuerdo del sujeto con cada una de ellas. Con este fin, los participantes emplearon una escala de 5 puntos, desde (1) *totalmente en desacuerdo* hasta (5) *totalmente de acuerdo*. Un ejemplo de ítem sería “*resisto las tentaciones con facilidad*”. La puntuación total de la escala se obtuvo mediante el sumatorio de las puntuaciones de cada ítem y el rango de puntuaciones puede oscilar entre 13 y 65. Puntuaciones altas indican mayor autocontrol.

El 50.33% de la muestra respondieron a esta escala. La Tabla 44 muestra la fiabilidad para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad: dos coeficientes por consistencia interna (Alfa de Cronbach y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente de fiabilidad test-retest y el coeficiente *CES*. Asimismo, la Tabla 44 recoge el valor del coeficiente de restricción en el rango (*U* y *u*) calculado para la muestra total del estudio. En este caso, no se pudieron calcular los coeficientes de restricción en el rango del grupo de hombres y mujeres por desconocer las desviaciones típicas de la población en ambos grupos.

Tabla 44.

Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango de la Escala Breve de Autocontrol

	α (n)	División 2 mitades (n)	Test-retest (n)	<i>CES</i> (n)	<i>U</i>	<i>u</i>
Muestra total	.84 (688)	.82 (688)	.81 (116)	.74 (116)	.98	1.02
Hombres	.84 (249)	.82 (249)	.60 (25)	.44 (25)		
Mujeres	.84 (249)	.83 (439)	.85 (91)	.81 (91)		

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; *CES* = coeficiente de estabilidad y equivalencia; *U* = coeficiente de restricción en el rango; $u = 1/U$.

Need for Cognition (NFC; Salgado, 2014b). Esta escala mide la disposición o tendencia de los individuos a buscar tareas que impliquen cognición y pensamiento y a disfrutar llevándolas a cabo (Cacioppo y Petty, 1982). La escala se compone de 10 ítems, donde los participantes tienen que indicar su grado de acuerdo con cada una de las afirmaciones. Para ello, se utiliza una escala de 5 puntos (desde 1 = *totalmente en desacuerdo* a 5 = *totalmente de acuerdo*). Un ejemplo de ítem sería “*evito la lectura de material difícil*”. La puntuación final de la escala se obtiene mediante el sumatorio de las puntuaciones de cada ítem. El rango de puntuaciones puede oscilar

entre 10 y 50, de modo que, puntuaciones altas en esta medida indican mayor disposición a la búsqueda y disfrute de tareas que impliquen esfuerzo cognitivo.

En este caso, el 61.16% de la muestra respondió a la medida de NFC. La Tabla 45 recoge la fiabilidad del instrumento para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad, dos por consistencia interna (Alfa de Cronbach y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente CES.

Tabla 45.

Coefficientes de Fiabilidad de la Escala NFC

	α (n)	División 2 mitades (n)	Test-retest (n)	CES (n)
Muestra total	.75 (836)	.60 (836)	.68 (129)	.24 (129)
Hombres	.73 (264)	.60 (264)	.49 (31)	.10 (31)
Mujeres	.75 (572)	.59 (572)	.71 (98)	.30 (98)

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; CES = coeficiente de estabilidad y equivalencia.

Test de Figuras Enmascaradas (TFE en su forma colectiva; Witkin, Oltman, Raskin y Karp, 1971). Se trata de un test perceptivo donde se presenta un conjunto de figuras complejas. La tarea de los participantes es tratar de localizar una figura simple, previamente mostrada, dentro de la figura grande y compleja, la cual ha sido diseñada de modo que oculte o enmascare la figura simple. Por lo tanto, el TFE permite evaluar la capacidad de “mantener los elementos separados de la experiencia”, es decir, mide el estilo perceptivo de dependencia-independencia de campo (Witkin et al., 1971). Los sujetos que presentan un estilo perceptivo dependiente de campo perciben los objetos dejándose influenciar por la organización general del campo que los rodea y los componentes son percibidos de forma difusa. Por su parte, quienes presentan un estilo perceptivo independiente son capaces de percibir los diversos componentes del campo de forma individual, pero manteniéndolos organizados dentro de su conjunto.

El TFE se compone de 3 secciones. La primera recoge 7 figuras complejas que deben ser resueltas en 2 minutos. La segunda y tercera sección presentan 9 ejercicios cada una de ellas y los sujetos disponen de 5 minutos para completar cada sección. El resultado de la corrección de la escala es el número total de figuras simples identificadas correctamente en la segunda y tercera sección. La primera sección se emplea para que los participantes practiquen y demuestren haber comprendido las instrucciones de la tarea. El rango de posibles puntuaciones oscila entre 0 y 18.

De este modo, puntuaciones altas en el TFE reflejan un estilo perceptivo más independiente de campo.

El TFE fue administrado al 16.83% de la muestra. La Tabla 46 presenta la fiabilidad de la medida para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad: dos coeficientes por consistencia interna (Kuder-Richardson 20 y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente *CES*. Asimismo, la tabla refleja los coeficientes de restricción en el rango (*U* y *u*) del test para la muestra del estudio.

Tabla 46.

Coefficientes de Fiabilidad y Coeficientes de Restricción en el Rango del Test de Figuras Enmascaradas

	<i>KR20</i> (<i>n</i>)	División 2 mitades (<i>n</i>)	Test-retest (<i>n</i>)	<i>CES</i> (<i>n</i>)	<i>U</i>	<i>u</i>
Muestra total	.86 (230)	.88 (230)	.84 (96)	.73 (96)	1.06	.94
Hombres	.85 (63)	.89 (63)	.80 (25)	.73 (25)	.95	1.05
Mujeres	.85 (167)	.86 (167)	.85 (71)	.73 (71)	1.07	.94

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; *KR20* = coeficiente de fiabilidad por consistencia interna Kuder-Richardson 20; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; *CES* = coeficiente de estabilidad y equivalencia; *U* = coeficiente de restricción en el rango; *u* = $1/U$.

Escala de Conductas Personales (CP5; Salgado, 2014d). Mide la disposición de los sujetos a actuar o comportarse de una determinada forma. Esta medida se divide en 5 subdimensiones que evalúan 5 disposiciones diferentes. La primera se denominada *impulsividad* y mide la tendencia de los sujetos a dar respuestas rápidas, tomar decisiones inmediatas o actuar, en general, de forma impulsiva. La segunda dimensión mide la *atención*, es decir, el grado de concentración o distracción en la realización de tareas y obligaciones. La tercera disposición que se mide es el *esfuerzo*, entendido como la dedicación, la persistencia en la tarea o trabajo a realizar y la fuerza de voluntad por acabar el trabajo. La cuarta disposición, denominada *intuición*, hace referencia al comportamiento espontáneo o no deliberado y a actuar de acuerdo con el instinto. Por último, la quinta subdimensión de la escala, *riesgo*, hace referencia al modo consistente de actuar de los sujetos ante situaciones inciertas o que impliquen riesgo, al grado de precaución adoptado en las distintas facetas de su vida y especialmente en lo relacionado con el futuro.

Cada una de las subdimensiones está compuesta por 12 ítems, en los que los participantes deben indicar su grado de acuerdo con cada una de las afirmaciones empleando una escala de 5 puntos (desde (1) *totalmente en desacuerdo* hasta (5) *totalmente de acuerdo*). Un ejemplo de ítem de cada una de las subescalas sería: (a) “*generalmente digo y hago las cosas con rapidez*”

(impulsividad), (b) “*me resulta difícil continuar con lo que estaba haciendo si me interrumpen a menudo*” (atención), (c) “*rara vez dejo las cosas a medias*” (esfuerzo), (d) “*suelo ser una persona muy intuitiva*” (intuición) y (e) “*no soy una persona propensa a correr ningún tipo de riesgo*” (riesgo). La puntuación individual en las subescalas se obtiene mediante el sumatorio de las puntuaciones de los ítems que las componen. El rango de puntuaciones puede oscilar entre 12 y 60 para cada una de las subdimensiones. Puntuaciones altas en las escalas implican mayor tendencia hacia cada uno de esos estilos cognitivos.

La escala CP5 fue administrada al 61.16% de la muestra. La Tabla 47 muestra la fiabilidad de las medidas para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad: dos coeficientes por consistencia interna (Alfa de Cronbach y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente CES.

Tabla 47.

Coefficientes de Fiabilidad de la Escala de Conductas Personales (CP5)

	α (n)	División 2 mitades (n)	Test-retest (n)	CES (n)
<i>Muestra total</i>				
Impulsividad	.63 (836)	.63 (836)	.70 (129)	.36 (129)
Atención	.71 (836)	.66 (836)	.53 (129)	.34 (129)
Esfuerzo	.81 (836)	.82 (836)	.73 (129)	.64 (129)
Intuición	.77 (836)	.83 (836)	.64 (129)	.56 (129)
Riesgo	.66 (836)	.69 (836)	.60 (129)	.46 (129)
<i>Hombres</i>				
Impulsividad	.62 (264)	.60 (264)	.38 (31)	.13 (31)
Atención	.70 (264)	.66 (264)	.54 (31)	.47 (31)
Esfuerzo	.79 (264)	.80 (264)	.51 (31)	.38 (31)
Intuición	.73 (264)	.80 (264)	.33 (31)	.24 (31)
Riesgo	.67 (264)	.70 (264)	.50 (31)	.39 (31)
<i>Mujeres</i>				
Impulsividad	.64 (572)	.64 (572)	.76 (98)	.57 (98)
Atención	.71 (572)	.66 (572)	.53 (98)	.30 (98)
Esfuerzo	.81 (572)	.83 (572)	.78 (98)	.70 (98)
Intuición	.79 (572)	.84 (572)	.69 (98)	.62 (98)
Riesgo	.65 (572)	.69 (572)	.62 (98)	.47 (98)

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; CES = coeficiente de estabilidad y equivalencia.

Desempeño

Se emplearon diferentes medidas de desempeño en este estudio. Para medir el desempeño académico se administraron diversas medidas tendientes a evaluar cada una de las dimensiones del desempeño. Así, para determinar el desempeño de tarea, se utilizó un instrumento de autoinforme, denominado CDTE (Salgado, 2010), y tres medidas objetivas: el GPA (*Grade Point Average*), las calificaciones alcanzadas en Bachillerato y la nota de acceso a la universidad (en adelante, CAU). Para evaluar el desempeño contextual se empleó la escala de autoinforme CDCE de Salgado (2010) y para determinar el desempeño contraproductivo se utilizó la escala de autoinforme CDAN (Salgado, 2010).

Para medir el desempeño ocupacional se administró un test de muestra de trabajo que evalúa el desempeño mediante una tarea administrativa denominada GCT o *General Clerical Test* (The Psychological Corporation Ltd., 1974). Seguidamente, sintetizamos las principales características de cada una de ellas.

Desempeño Académico

CDTE (Salgado, 2010). Es un instrumento de autoinforme compuesto por 30 ítems que evalúa el desempeño de tarea mediante 3 dimensiones: *cumplimiento de las obligaciones* (13 ítems), *orientación al logro* (8 ítems) e *implicación* (9 ítems). Para responder la escala, los sujetos deben indicar la frecuencia con la que llevan a cabo las conductas y acciones descritas en cada uno de los enunciados, utilizando una escala de 5 puntos (desde (1) *nunca* hasta (5) *siempre*). Un ejemplo de ítem sería: “*tomo notas en clase y hago apuntes en todas las asignaturas*”. La puntuación final de los individuos en la escala se obtiene mediante el sumatorio de las puntuaciones de cada ítem. El rango de puntuaciones puede oscilar entre 30 y 150, de modo que, puntuaciones altas en esta medida indican mejor desempeño de tarea.

La escala CDTE fue administrada al 69.86% de la muestra. La Tabla 48 muestra la fiabilidad de la medida para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad: dos por consistencia interna (Alfa de Cronbach y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente CES.

Tabla 48.

Coefficientes de Fiabilidad de la Escala CDTE

	α (n)	División 2 mitades (n)	Test-retest (n)	CES (n)
Muestra total	.89 (955)	.91 (955)	.79 (277)	.75 (277)
Hombres	.90 (311)	.92 (311)	.85 (80)	.79 (80)
Mujeres	.88 (644)	.90 (644)	.74 (197)	.70 (197)

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; CES = coeficiente de estabilidad y equivalencia.

GPA (Grade Point Average). Se refiere a las calificaciones promedio obtenidas en la titulación universitaria hasta el momento de realización del estudio. Para ello los participantes tuvieron que entregar una copia de su expediente académico, en el que se recogiese la nota media en base a 10. La calificación promedio de la muestra fue de 7.01 ($SD = .88$), en el grupo de hombres fue de 6.97 ($SD = .92$) y en el grupo de mujeres 7.03 ($SD = .87$).

Calificación de Bachillerato y Calificación de Acceso a la Universidad (CAU). Los estudiantes que participaron en este estudio entregaron una copia de la tarjeta de calificaciones de la EBAU con la nota media de acceso a la universidad. La EBAU es la prueba de acceso a la universidad española y consta de dos partes: una parte obligatoria y una parte voluntaria. Debido a que esta última parte es voluntaria no siempre se tiene en cuenta para calcular la nota de acceso a la universidad. La nota media de acceso a la universidad se conforma por el 60% de la calificación promedio de bachillerato y el 40% de la calificación promedio de la parte obligatoria de la EBAU. Asimismo, en los casos en que haya una evaluación de la parte voluntaria, ésta se añade a la nota de acceso a la universidad. Con el objetivo de conseguir una medida uniforme, la calificación de acceso a la universidad (en adelante CAU) empleada en nuestro estudio ha sido calculada de acuerdo con la siguiente fórmula: $CAU = .60 * \text{Nota media bachillerato} + .40 * \text{Calificación de la parte obligatoria de la prueba EBAU}$. De modo que, en los casos en que hubiese una calificación de la parte voluntaria, no se ha tenido en cuenta.

La calificación promedio de Bachillerato y de la CAU para la muestra del estudio fue de 7.67 ($SD = 1.22$) y 7.19 ($SD = 1.20$), respectivamente. Asimismo, los hombres alcanzaron una calificación promedio en Bachillerato de 7.71 ($SD = 1.29$) y en la CAU de 7.21 ($SD = 1.20$). Por su parte, las mujeres alcanzaron una calificación promedio en Bachillerato de 7.64 ($SD = 1.19$) y en la CAU de 7.18 ($SD = 1.21$).

CDCE (Salgado, 2010). Es un instrumento de autoinforme que mide el desempeño contextual. Está formada por 30 ítems y evalúa 3 dimensiones: *apoyo personal*, medida a través de 12 ítems, *apoyo organizacional* e *iniciativa consciente*, evaluadas con 9 ítems cada una de

ellas. Los sujetos deben indicar la frecuencia con la que llevan a cabo las conductas y acciones descritas en cada uno de los enunciados. Para ello deben emplear una escala de 1 (*nunca*) a 5 puntos (*siempre*). Un ejemplo de ítem de esta escala sería: “*ayudo a otros estudiantes cuando tienen dificultades con la materia*”. La puntuación final de los individuos en la escala se obtiene mediante el sumatorio de las puntuaciones de cada ítem. El rango de puntuaciones puede oscilar entre 30 y 150, de modo que, puntuaciones altas en esta medida indican mejor desempeño contextual.

La escala CDCE fue administrada al 69.35% de la muestra. La Tabla 49 recoge la fiabilidad de la medida para la muestra conjunta y para el grupo de hombres y mujeres. Al igual que en otros casos, se calcularon 4 coeficientes de fiabilidad: dos por consistencia interna (Alfa de Cronbach y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente CES.

Tabla 49.

Coefficientes de Fiabilidad de la Escala CDCE

	α (n)	División 2 mitades (n)	Test-retest (n)	CES (n)
Muestra total	.86 (948)	.91 (948)	.76 (215)	.71 (215)
Hombres	.87 (309)	.91 (309)	.83 (66)	.77 (66)
Mujeres	.86 (639)	.91 (639)	.73 (149)	.68 (149)

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; CES = coeficiente de estabilidad y equivalencia.

CDAN (Salgado, 2010). Es una medida de autoinforme que mide las conductas contraproductivas. Para ello, utiliza un conjunto de 30 ítems divididos en grupos de 6 enunciados en función de las diferentes dimensiones evaluadas. Concretamente, examina comportamientos y acciones en el ámbito académico considerados deshonestos, tales como copiar, uso inadecuado de recursos, absentismo, incumplimiento de normas y bajo esfuerzo consciente. Para responder a la escala, los sujetos deben indicar la frecuencia con la que llevan a cabo las conductas y acciones descritas en cada uno de los ítems, empleando una escala de 5 puntos (*nunca*, *casi nunca*, *a veces*, *casi siempre* y *siempre*). Un ejemplo de ítem sería: “*durante los exámenes he echado una ojeada a las respuestas de otro estudiante y usé esa información*”. La puntuación final de los individuos se obtiene mediante el sumatorio de las puntuaciones de cada ítem. El rango de puntuaciones puede oscilar entre 30 y 150, de modo que, puntuaciones altas en esta medida indican la comisión de más conductas contraproductivas.

La escala CDAN fue administrada al 64.89% de la muestra. La Tabla 50 recoge la fiabilidad de la medida para la muestra conjunta y para el grupo de hombres y mujeres. En este caso, también se calcularon 4 coeficientes de fiabilidad: dos por consistencia interna (Alfa de Cronbach y división en dos mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente *CES*.

Tabla 50.

Coefficientes de Fiabilidad de la Escala CDAN

	α (n)	División 2 mitades (n)	Test-retest (n)	<i>CES</i> (n)
Muestra total	.90 (887)	.93 (887)	.77 (274)	.73 (274)
Hombres	.91 (283)	.94 (283)	.79 (80)	.77 (80)
Mujeres	.89 (604)	.93 (604)	.74 (194)	.70 (194)

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; α = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; *CES* = coeficiente de estabilidad y equivalencia.

Desempeño Ocupacional

Test 2 del General Clerical Test (GCT; The Psychological Corporation Ltd.,1974). La batería GCT mide las aptitudes para lograr un desempeño administrativo eficaz y consta inicialmente de 9 tests. Algunos de ellos son de carácter aptitudinal y otros, como el Test 2, representan muestras de tareas administrativas. Los tests de la batería GCT se agrupan en: (1) aptitudes administrativas (GCT-C) que implican tareas de comprobación y clasificación; (2) aptitudes numéricas (GCT-N) que engloban tareas de cálculo aritmético, localización de errores y razonamiento aritmético y (3) aptitudes verbales (GCT-V) que recoge pruebas sobre comprensión lectora y relacionadas con la ortografía, el vocabulario y la sintaxis.

El Test 2, empleado en este estudio, forma parte de las pruebas englobadas dentro del grupo GCT-C o aptitudes administrativas y consiste en clasificar un listado de 61 archivos etiquetados por sus nombres en diferentes buzones. Los participantes disponen de 3 minutos para completar esta tarea. La puntuación final se obtiene mediante el sumatorio de archivos clasificados correctamente. El rango de puntuaciones puede oscilar entre 0 y 61, de modo que, puntuaciones altas en esta medida indican mejor desempeño de tarea.

Este test fue administrado al 34.75% de la muestra. La Tabla 51 recoge la fiabilidad de la medida para la muestra conjunta y para el grupo de hombres y mujeres. Se calcularon 4 coeficientes de fiabilidad: dos por consistencia interna (Kuder-Richardson 20 y división en dos

mitades con corrección por la fórmula de Spearman-Brown), el coeficiente test-retest y el coeficiente *CES*.

Tabla 51.

Coefficientes de Fiabilidad del Test 2 de la Batería GCT

	<i>KR20</i> (<i>n</i>)	División 2 mitades (<i>n</i>)	Test-retest (<i>n</i>)	<i>CES</i> (<i>n</i>)
Muestra Total	.96 (475)	.97 (475)	.79 (131)	.78 (131)
Hombres	.96 (135)	.97 (135)	.88 (32)	.96 (32)
Mujeres	.95 (340)	.97 (340)	.75 (99)	.71 (99)

Nota. Los valores entre paréntesis se corresponden con los tamaños de la muestra; *KR20* = coeficiente de fiabilidad por consistencia interna Kuder-Richardson 20; División 2 mitades = coeficiente de fiabilidad por consistencia interna mediante división en dos mitades, con corrección por la fórmula de Spearman-Brown; Test-retest = coeficiente de estabilidad temporal; *CES* = coeficiente de estabilidad y equivalencia.

3. RESULTADOS

Análisis de Correlaciones

Los Apéndices I, J y K muestran las correlaciones observadas entre el TRC y las variables examinadas en la muestra total y en el grupo de hombres y mujeres, respectivamente. De izquierda a derecha, las tablas reflejan las medias y las desviaciones típicas de cada medida. Seguidamente aparecen las correlaciones observadas entre las diversas variables. Las correlaciones entre el TRC-3 y el resto de las variables son correlaciones policóricas calculadas mediante el software FACTOR (Ferrando y Lorenzo-Seva, 2017). Las restantes son correlaciones de Pearson calculadas mediante el paquete estadístico SPSS. Los valores recogidos en la diagonal de la matriz son los coeficientes de fiabilidad empleados para corregir las correlaciones por error de medida. En el caso de las escalas, el coeficiente de fiabilidad empleado ha sido el Alfa de Cronbach, mientras que para los tests se empleó el coeficiente de correlación *CES*, con la excepción de las medidas RLS y Factor G de Cattell para las que se ha utilizado *KR20*, por no disponer de los datos necesarios para calcular el coeficiente *CES*. Los coeficientes de fiabilidad de las tres medidas de RC utilizados para la corrección por atenuación han sido los calculados en el estudio 1. Para el TRC-13 y TRC-10 se empleó el coeficiente *CES*. En el caso del TRC-3 se empleó el coeficiente test-retest. Asimismo, para aplicar las correcciones por error de medida de la variable criterio GPA, se empleó la distribución de la fiabilidad estimada por Cuadrado (2018). Esta distribución fue calculada mediante la integración de 7 coeficientes de fiabilidad Alfa de Cronbach y el resultado promedio obtenido ha sido .87 (*SD* = .06). Por último, en aquellos casos en los que se ha podido demostrar que las puntuaciones de las variables estaban restringidas en el rango (facilitación en el rango), se ha procedido a corregir los coeficientes de fiabilidad para evitar la

subestimación (sobreestimación) de sus magnitudes. Los valores reportados en las tablas son los valores ya corregidos por este fenómeno según cada caso. Para ello se ha empleado la fórmula del Caso I de Thorndike (1949).

Las Tablas 52, 53 y 54 muestran las correlaciones corregidas por atenuación de las medidas, para cada una de las variables estudiadas. De izquierda a derecha, se presenta, en primer lugar, la validez operativa, esto es, la correlación observada corregida por atenuación de la medida del criterio y restricción indirecta en el rango de la variable predictora. Seguidamente, se muestran los intervalos de confianza del 95% de la validez operativa. La cuarta columna refleja las correlaciones verdaderas o, lo que es lo mismo, las correlaciones observadas corregidas por todos los artefactos de error en la medida del predictor y criterio y restricción indirecta en el rango del predictor. Finalmente, la quinta columna recoge los intervalos de confianza del 95% de la validez verdadera.

En los casos en que se pudo calcular el coeficiente de restricción en el rango, las correlaciones también fueron corregidas por restricción indirecta en el rango de la variable predictora. En los casos particulares en que no se haya podido calcular el coeficiente de restricción en el rango para las muestras de hombres y mujeres (véase, por ejemplo, Factor G y la escala de autocontrol), se ha empleado el coeficiente calculado para la muestra general, al tratarse éste de una buena aproximación. La fórmula empleada para corregir las correlaciones por error en las medidas fue la fórmula clásica de corrección por atenuación (véase, por ejemplo, Hunter y Schmidt, 2015). Las correcciones por restricción indirecta en el predictor fueron halladas mediante la fórmula del Caso II de Thorndike (1949; véase también Hunter y Schmidt, 2015). Para llevar a cabo estas correcciones se empleó el software VALCOR de Salgado (1997).

Validez Convergente y Discriminante

Capacidad Mental General (CMG) y Reflexibilidad Cognitiva (RC)

La Tabla 52 muestra las correlaciones corregidas entre el TRC y las diferentes medidas de CMG para la muestra total y los grupos de hombres y mujeres. Como se puede observar, el TRC-13, TRC-10 y TRC-3 se asociaron con la CMG con independencia de la muestra y medida de capacidad cognitiva. En todos los casos, las correlaciones verdaderas y operativas fueron significativas, ya que los intervalos de confianza del 95% no incluyeron el cero. De acuerdo con los criterios de Pearson (1904), la RC se relacionó de forma considerable con las tres medidas de capacidad cognitiva en la muestra total. Los tamaños del efecto verdadero oscilaron entre .644 y .769 en el WPT, entre .477 y .596 en el Factor G de Cattell y entre .509 y .538 en el RLS. Las correlaciones operativas oscilaron entre .601 y .722 en el WPT, entre .403 y .502 en el Factor G de Cattell y entre .423 y .448 en el RLS. De este modo, los resultados apoyan la Hipótesis 1.

Además, los tamaños del efecto operativos y verdaderos entre el WPT y las tres medidas de RC fueron más robustos que con el Factor G y con el RLS, siendo estas diferencias estadísticamente significativas (rango de $z = 5.53$ a 7.29 en el TRC-13; 5.26 a 6.63 en el TRC-10 y 3.14 a 4.25 en el TRC-3). No obstante, las diferencias entre los tamaños del efecto de los diferentes tipos de TRC y el Factor G de Cattell y RLS no fueron significativas. El WPT es un instrumento que se apoya en la capacidad numérica en mayor proporción que las otras dos medidas de CMG. Este dato podría sugerir que las diferencias en la relación entre el TRC y las medidas de CMG podrían venir determinadas por la mayor capacidad numérica que exigen el TRC y el WPT. En cualquier caso, estos resultados reflejan que la medida WPT se relaciona de forma más robusta con la RC que el Factor G de Cattell o el test RLS.

Los resultados también han mostrado que los tamaños del efecto entre el TRC-3 y las restantes medidas de CMG fueron más pequeños que los hallados con el TRC-13 y TRC-10. No obstante, solo se encontraron diferencias significativas entre los tamaños del efecto verdadero entre el WPT y TRC-3 y TRC-10 ($z = 2.15$).

En esta misma línea se encuentran los resultados hallados en la muestra de hombres y mujeres. En ambos casos, la RC se relacionó de forma considerable con las tres medidas de CMG. En el grupo de los hombres, las correlaciones corregidas oscilaron entre $.617$ a $.727$ en el WPT, entre $.569$ y $.716$ en el Factor G de Cattell y entre $.324$ y $.505$ en el RLS. En el grupo de mujeres, las correlaciones corregidas oscilaron entre $.615$ y $.774$ en el WPT, entre $.493$ y $.770$ en el Factor G de Cattell y entre $.443$ y $.571$ en el RLS. Además, los tamaños del efecto entre las tres medidas de RC y el WPT fueron más robustos que con el Factor G o el RLS. No obstante, en el caso de los hombres estas diferencias únicamente fueron significativas entre las correlaciones con WPT y RLS. Estos resultados parecen sugerir que el WPT podría explicar más proporción de varianza de la RC que el resto de las medidas de capacidad cognitiva examinadas.

Asimismo, los resultados de hombres y mujeres también han reflejado que los tamaños del efecto entre el TRC-3 y las medidas de CMG fueron más pequeños que los hallados entre el TRC-13 y el TRC-10 con esas variables. No obstante, en el grupo de los hombres estas diferencias únicamente no fueron significativas para el caso del TRC-13, donde todas las medidas de CMG explicaron la misma proporción de RC. Sin embargo, para el grupo de mujeres únicamente las diferencias entre las correlaciones con RLS no fueron significativas, lo que indica que tanto el WPT como el Factor G explican más varianza de la RC que el RLS, con independencia del instrumento de RC empleado.

Tabla 52.

Correlaciones Corregidas entre el TRC y las Medidas de CMG

	TRC-13						TRC-10						TRC-3					
	IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)		
	r_{op}	LI	LS	ρ	LI	LS	r_{op}	LI	LS	ρ	LI	LS	r_{op}	LI	LS	ρ	LI	LS
<i>N total</i>																		
WPT	.700	.664	.736	.746	.715	.777	.722	.689	.756	.769	.741	.797	.601	.554	.648	.644	.602	.686
Factor G	.477	.395	.560	.567	.494	.639	.502	.422	.582	.596	.527	.665	.403	.313	.492	.477	.395	.560
RLS	.447	.354	.539	.536	.455	.618	.448	.356	.540	.538	.456	.619	.423	.329	.518	.509	.424	.594
<i>Hombres</i>																		
WPT	.645	.573	.717	.690	.627	.754	.681	.615	.746	.727	.671	.784	.574	.490	.659	.617	.540	.694
Factor G	.589	.481	.697	.675	.585	.765	.624	.523	.725	.716	.635	.797	.496	.372	.621	.569	.457	.681
RLS	.345	.160	.531	.398	.222	.574	.281	.085	.476	.324	.135	.513	.440	.272	.608	.505	.351	.659
<i>Mujeres</i>																		
WPT	.699	.655	.744	.747	.709	.784	.726	.685	.766	.774	.740	.807	.572	.511	.633	.615	.559	.671
Factor G	.566	.470	.662	.703	.631	.774	.620	.533	.707	.770	.712	.827	.398	.279	.516	.493	.387	.600
RLS	.410	.301	.519	.541	.447	.635	.432	.325	.539	.571	.481	.661	.337	.222	.453	.443	.337	.549

Nota. TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); TRC-10 = ítems Salgado (2014a); TRC-3 = ítems Frederick (2005); r_{op} = correlación operativa o correlación corregida por error de medida en el criterio y restricción indirecta en el rango en el predictor; IC 95% (r_{op}) = intervalos de confianza del 95% de la correlación operativa; ρ = correlación verdadera o correlación observada corregida por error de la medida del predictor y criterio y por restricción indirecta en el rango del predictor; IC 95% (ρ) = intervalos de confianza del 95% de la correlación verdadera; LI = límite inferior; LS = límite superior; N total = muestra total; WPT = Wonderlic personnel test; RLS = test de razonamiento lógico simbólico.

Finalmente, si comparamos las correlaciones halladas entre las medidas de RC y las de CMG en el grupo de hombres y de mujeres, podemos observar que no existen diferencias significativas en cuanto a la capacidad de esas medidas para determinar el desempeño del TRC en hombres y mujeres.

Personalidad y Reflexibilidad Cognitiva (RC)

Las correlaciones corregidas entre los 5 factores de personalidad y las medidas de RC se presentan en la Tabla 53. Para cada una de las muestras, se presentan, en primer lugar, los resultados empleando la medida de personalidad IP/5F y, seguidamente, los resultados correspondientes a la medida de personalidad quasi-ipsativa, QI5F-Tri.

En la muestra general, los resultados indican que los factores de personalidad de EE y AP, medidos mediante el IP/5F, fueron las únicas variables de personalidad que se relacionaron con el desempeño en las tres medidas de RC. En el caso de EE, los tamaños del efecto operativos fueron .184, .187 y .158 para el TRC-13, TRC-10 y TRC-3, respectivamente; mientras que los tamaños del efecto verdaderos fueron .193, .195 y .166. En el caso de AP, los tamaños del efecto fueron muy similares a los de EE, aunque ligeramente inferiores. Así, por ejemplo, los tamaños del efecto operativos fueron .142, .129 y .144 para el TRC-13, TRC-10 y TRC-3, respectivamente y los tamaños del efecto teóricos fueron .155, .142 y .157.

Respecto a la medida de personalidad QI5F-Tri, los resultados mostraron que el factor AP también se relacionó con los resultados en las tres medidas de RC, con correlaciones teóricas de .210, .205 y .203 para el TRC-13, TRC-10 y TRC-3, respectivamente. Además, aunque los tamaños del efecto fueron más robustos con esta medida de personalidad, las diferencias no fueron significativas. Asimismo, los resultados mostraron que el factor C también se relacionó con la RC. No obstante, la correlación encontrada fue negativa y pequeña ($p = -.125$ para el TRC-13; $p = -.127$ para el TRC-10 y $p = -.118$ para el TRC-3).

Los resultados encontrados para la muestra de hombres y de mujeres también reflejaron que el factor AP correlacionó con la RC con independencia de la medida de personalidad y RC empleada. Las correlaciones verdaderas en el grupo de los hombres oscilaron entre .145 y .146, en la medida IP/5F, y entre .150 y .177, en la medida QI5F-Tri. En el grupo de las mujeres, las correlaciones teóricas oscilaron entre .114 a .147, en la medida IP/5F, y entre .201 y .211, en la medida QI5F-Tri. En ningún caso se encontraron diferencias significativas entre los tamaños del efecto de la RC y AP en función del instrumento de personalidad o RC. Tampoco se encontraron diferencias significativas entre hombres y mujeres en cuanto a la relación de la RC y el factor AP.

Tabla 53.

Correlaciones Corregidas entre el TRC y las Medidas de Personalidad

	TRC-I3						TRC-I0						TRC-3					
	IC 95% (r _{op})			IC 95% (p)			IC 95% (r _{op})			IC 95% (p)			IC 95% (r _{op})			IC 95% (p)		
	r _{op}	LI	LS	ρ	LI	LS	r _{op}	LI	LS	ρ	LI	LS	r _{op}	LI	LS	ρ	LI	LS
<i>N total</i>																		
IP/5F																		
EE	.184	.104	.263	.193	.113	.272	.187	.107	.266	.195	.116	.274	.158	.078	.238	.166	.086	.246
EX	.054	-.033	.141	.059	-.028	.146	.042	-.045	.129	.046	-.041	.132	.081	-.005	.168	.089	.002	.175
AP	.142	.065	.218	.155	.079	.231	.129	.053	.206	.142	.065	.218	.144	.067	.220	.157	.081	.233
AM	-.028	-.098	.042	-.031	-.101	.038	-.044	-.114	.025	-.050	-.120	.019	.011	-.059	.081	.012	-.058	.082
C	-.071	-.152	.009	-.077	-.157	.003	-.073	-.154	.007	-.080	-.160	.000	-.066	-.147	.014	-.072	-.152	.009
Q15F-Tri																		
EE	.055	-.024	.135	.068	.002	.134	.083	-.003	.170	.103	.038	.168	-.009	-.087	.068	-.012	-.078	.054
EX	-.037	-.117	.043	-.043	-.109	.023	-.029	-.116	.057	-.034	-.100	.032	-.053	-.130	.023	-.062	-.128	.004
AP	.192	.110	.274	.210	.147	.273	.187	.097	.277	.205	.142	.268	.185	.107	.263	.203	.140	.266
AM	-.026	-.106	.054	-.031	-.097	.035	-.060	-.146	.027	-.072	-.137	.007	.055	-.022	.131	.066	.000	.132
C	-.110	-.189	-.030	-.125	-.190	-.060	-.111	-.198	-.025	-.127	-.192	-.062	-.103	-.180	-.027	-.118	-.183	-.053
<i>Hombres</i>																		
IP/5F																		
EE	.095	-.035	.226	.101	-.029	.231	.101	-.030	.231	.107	-.023	.237	.075	-.056	.206	.079	-.052	.210
EX	.061	-.080	.201	.066	-.075	.206	.016	-.125	.157	.017	-.124	.159	.114	-.025	.254	.125	-.014	.264
AP	.135	.016	.253	.146	.028	.265	.134	.015	.253	.145	.027	.264	.133	.014	.252	.145	.026	.263
AM	-.126	-.228	-.023	-.140	-.242	-.038	-.204	-.304	-.104	-.228	-.327	-.129	.019	-.084	.123	.021	-.082	.125
C	-.120	-.252	.011	-.131	-.262	.000	-.094	-.227	.038	-.103	-.235	.029	-.159	-.289	-.029	-.173	-.302	-.043
Q15F-Tri																		
EE	.088	-.057	.233	.109	-.010	.228	.146	-.025	.316	.181	.064	.298	-.017	-.155	.120	-.021	-.142	.100
EX	.078	-.067	.223	.090	-.030	.210	.104	-.063	.271	.120	.001	.239	.015	-.123	.152	.017	-.104	.138
AP	.160	.014	.307	.174	.057	.291	.162	-.010	.334	.177	.060	.294	.138	.003	.273	.150	.032	.268
AM	-.044	-.189	.101	-.052	-.172	.068	-.067	-.232	.099	-.078	-.198	.042	.000	.138	.138	.000	-.138	.138
C	-.241	-.393	-.089	-.278	-.389	-.167	-.262	-.449	-.075	-.303	-.413	-.193	-.200	-.336	-.063	-.231	-.345	-.117

Continúa

Tabla 53.
Continuación

	TRC-13						TRC-10						TRC-3					
	IC 95% (τ_{op})			IC 95% (ρ)			IC 95% (τ_{op})			IC 95% (ρ)			IC 95% (τ_{op})			IC 95% (ρ)		
	τ_{op}	LI	LS	ρ	LI	LS	τ_{op}	LI	LS	ρ	LI	LS	τ_{op}	LI	LS	ρ	LI	LS
<i>Mujeres</i>																		
IP/5F																		
EE	-.019	-.125	.086	-.020	-.126	.085	-.031	-.136	.074	-.032	-.138	.073	-.006	-.112	.099	-.006	-.112	.099
EX	.000	-.111	.111	.000	-.111	.111	.000	-.111	.111	.000	-.111	.111	.008	-.103	.119	.009	-.102	.120
AP	.123	.022	.224	.133	.033	.234	.104	.003	.205	.114	.012	.215	.134	.033	.235	.147	.046	.247
AM	.048	-.044	.139	.055	-.037	.146	.066	-.025	.158	.076	-.015	.168	.010	-.082	.102	.012	-.080	.104
C	.004	-.097	.105	.004	-.096	.105	-.014	-.115	.086	-.015	-.116	.085	.034	-.067	.134	.036	-.065	.136
<i>Q15F-Tri</i>																		
EE	-.076	-.175	.024	-.095	-.173	-.017	-.062	-.169	.045	-.078	-.156	.000	-.100	-.193	-.007	-.125	-.202	-.048
EX	-.105	-.204	-.006	-.121	-.198	-.044	-.102	-.209	.006	-.117	-.194	-.040	-.100	-.193	-.007	-.115	-.193	-.037
AP	.192	.089	.295	.211	.136	.286	.183	.071	.295	.201	.126	.276	.188	.093	.282	.206	.131	.281
AM	-.001	-.102	.099	-.002	-.081	.077	-.044	-.151	.063	-.054	-.132	.024	.094	.001	.187	.114	.036	.192
C	.009	-.091	.109	.010	-.069	.089	.014	-.094	.121	.016	-.063	.095	.000	-.095	.095	.000	-.095	.095

Nota. TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); TRC-10 = ítems Salgado (2014a); TRC-3 = ítems Frederick (2005); τ_{op} = correlación operativa o correlación corregida por error de medida en el criterio y restricción indirecta en el rango en el predictor; IC 95% (τ_{op}) = intervalos de confianza del 95% de la correlación operativa; ρ = correlación verdadera o correlación observada corregida por error de la medida del predictor y criterio y por restricción indirecta en el rango del predictor; IC 95% (ρ) = intervalos de confianza del 95% de la correlación verdadera; LI = límite inferior; LS = límite superior; EE = estabilidad emocional; EX = extraversión; AP = apertura a la experiencia; AM = amigabilidad; C = concuencia.

En el grupo de los hombres, el factor C, medido con el Q15F-Tri, y el factor de AM, medido con el IP/5F, también fueron determinantes de la RC. De acuerdo con los criterios de Pearson (1904), las correlaciones con el factor de C fueron moderadas (rango $r_{op} = -.200$ a $-.262$ y rango $\rho = -.231$ a $-.303$) y las correlaciones con AM fueron pequeñas ($r_{op} = -.126$ y $-.204$ y $\rho = -.140$ y $-.228$) y, en ambos casos, los tamaños del efecto operativos y teóricos fueron negativos.

En el grupo de las mujeres, los resultados mostraron que, además de AP, el factor de EX también determinó la RC. Los tamaños del efecto operativos y teóricos hallados fueron negativos y pequeños (rango $r_{op} = -.105$ a $-.100$ y rango $\rho = -.115$ a $-.121$). Por otro lado, los factores de AM y EE correlacionaron, únicamente, con el TRC-3 pero no con las otras dos medidas de RC. Además, esta correlación fue pequeña en ambos casos, lo que sugiere que la RC de las mujeres podría no estar determinada realmente por estos dos factores.

Por lo tanto, los resultados de este estudio apoyan nuestra Hipótesis 2a y rechazan parcialmente la Hipótesis 2b. Los datos mostraron que los factores de EE, AP (medidos con el IP/5F) y C (medido con el Q15F-Tri) determinan la RC del total de la muestra. No obstante, cuando se examinó la capacidad de los factores de personalidad para explicar la RC en hombres y mujeres, los resultados mostraron que el factor de personalidad AP determina la RC de ambos grupos, con independencia del inventario de personalidad y del instrumento de RC empleados. Asimismo, también se encontró que los factores de C y AM fueron determinantes de la RC en el grupo de los hombres y que el factor de EX lo fue en el grupo de las mujeres. Por lo que, los factores de personalidad que determinan la RC podrían ser diferentes en ambos sexos. El resto de los factores no mostraron relación con el constructo o mostraron una pequeña asociación con alguno de los instrumentos de RC, lo que podría sugerir que no son determinantes de este constructo.

Disposiciones de Pensamiento (DP) y Reflexibilidad Cognitiva (RC)

La Tabla 54 refleja las correlaciones operativas y teóricas entre el TRC y las medidas de DP examinadas. En primer lugar, se presentan las correlaciones corregidas para la muestra total seguidas de las correlaciones corregidas obtenidas para el grupo de hombres y de mujeres.

Como se puede observar, los resultados muestran que solo algunas medidas de DP explicaron el desempeño en el TRC. Así, por ejemplo, la disposición hacia un pensamiento comprometido y de disfrute con el esfuerzo cognitivo (NFC) se relacionó con el desempeño en el TRC, tanto en la muestra total, como en las muestras de hombres y de mujeres. De acuerdo con los criterios de Pearson (1904), los tamaños del efecto encontrados han sido moderados. Así, en el caso de la muestra total, las correlaciones verdaderas oscilaron entre .316 y .353; en la muestra de hombres, entre .307 y .402; y en el grupo de mujeres, entre .259 y .285. Los tamaños del efecto

fueron más robustos en el grupo de hombres que en el de las mujeres, pero en ningún caso las diferencias fueron significativas ($z < 1.96$).

Del mismo modo, los resultados mostraron que la RC se asoció de forma significativa con el estilo cognitivo de dependencia-independencia de campo (TFE). Además, esta asociación fue la más robusta encontrada en comparación con el resto de las medidas de DP. No obstante, estos resultados deben ser tomados con precaución ya que los tamaños del efecto se encuentran afectados, en mayor medida, por el error de muestreo. Así pues, los datos muestran correlaciones teóricas moderadas en la muestra total (rango de $\rho = .433$ a $.506$) y considerables en la muestra de hombres (rango de $\rho = .459$ a $.491$) y de mujeres (rango de $\rho = .426$ a $.608$). En este caso tampoco se encontraron diferencias significativas en la capacidad del TFE para determinar la RC en hombres y mujeres.

En relación con la escala CP5, los datos mostraron una relación entre la RC y las subescalas de atención y de riesgo en las tres muestras. Las correlaciones teóricas entre el TRC y el estilo cognitivo orientado a la atención fueron pequeñas, tanto en la muestra total (rango de $\rho = .140$ a $.194$) como en el grupo de hombres (rango de $\rho = .164$ a $.198$) y de mujeres (rango de $\rho = .097$ a $.192$). Estos resultados reflejan que la RC de los individuos podría estar determinada por el estilo cognitivo de atención en la tarea, ya que la disposición a mantenerse concentrado en la actividad y no distraerse de forma fácil podría ayudar a detectar más fácilmente la respuesta intuitiva y engañosa del TRC e iniciar una mayor deliberación en busca de la respuesta correcta. Además, las diferencias entre los tamaños del efecto de cada uno de los TRCs y la subescala de atención de hombres y mujeres no resultaron significativos, lo que sugiere que esta medida es un determinante del constructo igualmente válido en ambos grupos.

Los resultados también mostraron que la RC se asoció con la disposición a asumir riesgos. Los tamaños del efecto teóricos encontrados para la muestra total (rango de $\rho = .133$ y $.201$) y el grupo de hombres (rango de $\rho = .199$ a $.310$) y mujeres (rango de $\rho = .058$ y $.116$) fueron pequeños. En este último caso, la correlación con el TRC-3 fue prácticamente nula y no significativa. También se encontraron diferencias significativas entre los tamaños del efecto de hombres y mujeres entre riesgo y TRC-13 ($z = 2.45$) y TRC-10 ($z = 2.73$), pero no con el TRC-3 ($z = 1.91$). Estos datos sugieren que la disposición a asumir riesgos puede determinar en mayor medida el desempeño de hombres que el de mujeres en el TRC-13 y TRC-10. Finalmente, también se encontró que la DP de esfuerzo fue determinante de la RC de hombres (rango de $\rho = .045$ a $.136$) pero no de mujeres y que la DP de autocontrol lo fue de mujeres (rango de $\rho = -.080$ a $-.127$) pero no de hombres.

Tabla 54.

Correlaciones Corregidas entre el TRC y las Medidas de Disposición de Pensamiento

	TRC-13						TRC-10						TRC-3					
	IC 95% (r _{op})			IC 95% (ρ)			IC 95% (r _{op})			IC 95% (ρ)			IC 95% (r _{op})			IC 95% (ρ)		
	r _{op}	LI	LS	ρ	LI	LS	r _{op}	LI	LS	ρ	LI	LS	r _{op}	LI	LS	ρ	LI	LS
<i>N total</i>																		
Autocontrol	-.043	-.116	.030	-.047	-.120	.026	-.043	-.116	.030	-.047	-.119	.026	-.041	-.114	.032	-.044	-.117	.029
NFC	.303	.213	.393	.349	.289	.409	.306	.204	.408	.353	.294	.412	.274	.191	.357	.316	.255	.377
TFE	.435	.325	.545	.506	.406	.607	.432	.322	.542	.503	.402	.604	.372	.255	.489	.433	.323	.543
<i>CP5</i>																		
Impulsiv.	.021	-.062	.104	.026	-.042	.094	.001	-.088	.091	.002	-.066	.070	.061	-.019	.140	.076	.009	.143
Atención	.154	.071	.237	.183	.117	.249	.163	.072	.254	.194	.129	.259	.117	.039	.196	.140	.074	.206
Esfuerzo	.038	-.044	.121	.043	-.025	.111	.046	-.042	.135	.052	-.016	.120	.014	-.066	.094	.016	-.052	.084
Intuición	.042	-.041	.124	.048	-.020	.116	.049	-.040	.138	.056	-.012	.124	.024	-.056	.103	.027	-.041	.095
Riesgo	.149	.066	.232	.184	.118	.250	.163	.072	.254	.201	.136	.266	.108	.029	.187	.133	.066	.200
<i>Hombres</i>																		
Autocontrol	-.030	-.152	.091	-.032	-.154	.089	-.007	-.128	.115	-.007	-.129	.114	-.067	-.187	.054	-.072	-.193	.049
NFC	.314	.155	.473	.368	.264	.472	.343	.140	.545	.402	.301	.503	.262	.122	.401	.307	.198	.416
TFE	.413	.216	.610	.484	.302	.667	.392	.191	.593	.459	.271	.646	.419	.223	.615	.491	.310	.672
<i>CP5</i>																		
Impulsiv.	.060	-.085	.205	.076	-.044	.196	.062	-.103	.228	.079	-.041	.199	.042	-.094	.179	.054	-.066	.174
Atención	.144	-.002	.290	.172	.055	.289	.137	-.032	.307	.164	.047	.281	.166	.031	.302	.198	.082	.314
Esfuerzo	.095	-.049	.240	.107	-.012	.226	.121	-.048	.289	.136	.018	.254	.040	-.096	.177	.045	-.075	.165
Intuición	.075	-.070	.219	.087	-.033	.207	.073	-.092	.239	.086	-.034	.206	.071	-.064	.207	.083	-.037	.203
Riesgo	.226	.075	.377	.276	.165	.387	.254	.068	.439	.310	.201	.419	.163	.027	.298	.199	.083	.315

Continúa

Tabla 54.
Continuación

	TRC-13						TRC-10						TRC-3					
	IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)		
	r_{op}	LI	LS	ρ	LI	LS	r_{op}	LI	LS	ρ	LI	LS	r_{op}	LI	LS	ρ	LI	LS
<i>Mujeres</i>																		
Autocontrol	-.103	-.193	-.012	-.112	-.202	-.021	-.117	-.208	-.027	-.127	-.217	-.037	-.073	-.164	.018	-.080	-.170	.011
NFC	.247	.136	.358	.285	.210	.360	.242	.119	.365	.279	.203	.355	.225	.125	.325	.259	.183	.335
TFE	.501	.382	.621	.584	.479	.688	.523	.407	.638	.608	.508	.708	.365	.225	.504	.426	.295	.558
<i>CP5</i>																		
Impulsiv.	-.010	-.115	.094	-.013	-.095	.069	-.041	-.153	.070	-.052	-.134	.030	.058	-.039	.155	.072	.001	.143
Atención	.143	.039	.248	.170	.090	.250	.162	.047	.278	.192	.113	.271	.082	-.015	.179	.097	.016	.178
Esfuerzo	.004	-.101	.109	.004	-.078	.086	.008	-.104	.121	.009	-.073	.091	-.011	-.109	.087	-.012	-.094	.070
Intuición	.015	-.089	.120	.017	-.065	.099	.027	-.084	.139	.031	-.051	.113	-.012	-.110	.086	-.014	-.096	.068
Riesgo	.081	-.023	.184	.100	.019	.181	.093	-.019	.205	.116	.035	.197	.047	-.050	.144	.058	-.024	.140

Nota. TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); TRC-10 = ítems Salgado (2014a); TRC-3 = ítems Frederick (2005); r_{op} = correlación operativa o correlación corregida por error de medida en el criterio y restricción indirecta en el rango en el predictor; IC 95% (r_{op}) = intervalos de confianza del 95% de la correlación operativa; ρ = correlación verdadera o correlación observada corregida por error de la medida del predictor y criterio y por restricción indirecta en el rango del predictor; IC 95% (ρ) = intervalos de confianza del 95% de la correlación verdadera; LI = límite inferior; LS = límite superior; N total = muestra total; NFC = need for cognition; TFE = test de figuras enmascaradas; CP5 = escala de conductas personales; Impulsiv. = impulsividad.

En resumen, los resultados sobre la relación entre la RC y las DP sugieren que la disposición a asumir riesgos y a comprometerse con el esfuerzo cognitivo, y los estilos cognitivos de independencia de campo y atención y concentración en la tarea son determinantes del desempeño en el TRC. Estos resultados apoyan la Hipótesis 3. Además, los datos reflejan que estas medidas son igualmente válidas para explicar la RC de hombres y de mujeres. Únicamente se han encontrado diferencias significativas en la capacidad de la subescala de riesgo del CP5, donde los datos mostraron que explica de forma más robusta varianza del desempeño de los hombres que de las mujeres en el TRC-13 y TRC-10. Asimismo, la disposición a autocontrolar nuestros impulsos fue un determinante de la RC de las mujeres, pero no de los hombres, mientras que la DP de esfuerzo lo fue de los hombres, pero no de las mujeres. Estos datos sugieren que las variables que explican la RC en ambos grupos podrían ser diferentes. Finalmente, las disposiciones de pensamiento medidas con las restantes subescalas del CP5 (impulsividad, esfuerzo e intuición) no explicaron varianza de la RC.

Predicción Múltiple del Desempeño en el TRC-13

Una vez halladas las relaciones entre las variables individuales y la RC, se ha explorado la proporción de varianza de la RC que, en conjunto, son capaces de explicar estas variables y cuáles de ellas contribuyen de forma independiente en la determinación de la RC. Para ello, se ha llevado a cabo un análisis de regresión en tres pasos, empleando el TRC-13 como variable dependiente. Estudios previos han sugerido que el TRC es una medida principalmente de CMG y los resultados de este estudio así lo confirman. La variable individual que presenta una relación más robusta con la RC ha sido la CMG, y entre ellas, el WPT. Por lo tanto, (1) la primera variable introducida en el análisis de regresión ha sido la CMG. Como las tres medidas de capacidad cognitiva están muy correlacionadas entre sí, se utilizó únicamente el WPT. (2) En el segundo paso, hemos introducido los factores de personalidad, además de la CMG, ya que los datos de nuestro estudio demuestran que la personalidad y, en alguna ocasión, CMG son determinantes de las DP (véase Tabla 55). La medida de personalidad introducida ha sido el inventario QI5F-Tri, por ser el que presenta un mayor tamaño de la muestra. (3) En el tercer paso, se han añadido las variables de DP, de modo que el modelo predictivo se conforma de las variables CMG, personalidad y DP. La DP de dependencia-independencia de campo (medido con el TFE) no ha sido incorporada al análisis de regresión debido al tamaño reducido de la muestra. Este tercer paso ha sido repetido varias veces hasta depurar las correlaciones entre las variables, es decir, hasta eliminar todas las variables que no contribuyen a la determinación de la RC. De este modo, el análisis de regresión presenta un cuarto modelo, que es el modelo final, donde se recogen únicamente las variables que de forma independiente explican el desempeño en el TRC.

Tabla 55.

Determinantes de las Disposiciones de Pensamiento (Muestra Total)

	Impulsividad (β)	Intuición (β)	Riesgo (β)
WPT	-	-	-.06**
EE	-.47***	-.18***	-
EX	-.28***	-.10*	-
AP	-.36***	-.05	.12***
AM	-.74***	-.37***	-.15***
C	-1.00***	.60***	-.78***
<i>F</i>	417.80	86.03	454.81
<i>R</i>	.81***	.53***	.79***
<i>R</i> ²	.65	.28	.62
<i>R</i> _{CV}	.80	.52	.79
<i>R</i> ² _{CV}	.65	.27	.62

Nota. WPT = Wonderlic personnel test; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

p* < .05; *p* < .01; ****p* < .001

Para llevar a cabo los análisis de regresión se han empleado las correlaciones corregidas. Cuando se desea explorar la predicción múltiple de una determinada variable se recomienda, desde el punto de vista metodológico, utilizar las correlaciones corregidas ya que, en caso contrario, los parámetros de regresión estimados pueden estar sesgados y los resultados pueden llevar a inferencias erróneas (Fritz, Kenny y MacKinnon, 2016; Hunter y Schmidt, 2015). Siguiendo esta recomendación, la Tabla 56 presenta la matriz de correlaciones corregidas utilizada en el análisis de regresión de la muestra total.

Los cálculos se han ejecutado mediante el paquete estadístico SPSS y los resultados reportan los pesos estandarizados de regresión de cada una de las variables, así como la correlación múltiple (*R*) y el coeficiente de determinación múltiple corregido (*R*²). Se debe tener en cuenta que, cuando se pretende estimar el coeficiente de correlación múltiple de la población partiendo de una muestra determinada de sujetos, los resultados obtenidos suelen estar estadísticamente sesgados. El proceso de optimización a partir del cual se deriva la ecuación de regresión hace que *R* y, por consiguiente, *R*² sean sistémicamente más altos que los parámetros correspondientes a la población. Consecuentemente, cuando la ecuación de regresión se aplica a una muestra independiente, diferente de la que se obtiene la ecuación, el poder predictivo disminuye (Yin y Fan, 2001). Por esta razón, los coeficientes de la población *R* y *R*² que el programa estadístico SPSS ofrece no son los estimadores más adecuados de validez múltiple. Para

determinar la generalización o el poder predictivo de la correlación múltiple obtenida en esta muestra se ha estimado el coeficiente de correlación múltiple cruzado de la población (R_{CV}) y el coeficiente de determinación múltiple cruzado de la población (R^2_{CV}). Para ello, se ha empleado la fórmula de Browne (1975) y se han seguido los procedimientos recomendados por Lautenschlager (1990) y Yin y Fan (2001) en relación con el cálculo del estimador p^4 de Wherry (1931), incluido en la fórmula de Browne.

La Tabla 57 recoge los resultados hallados en el análisis de regresión de la muestra total. Como se puede observar, el modelo 4 es el modelo que explicó mayor proporción de varianza de la RC ($R^2_{CV} = .63$). El modelo incluye todas las variables objeto de estudio (CMG, personalidad y DP) y es el resultado de haber eliminado las variables del modelo 3 que no contribuyen a la predicción de la RC (EX, AP, AM, autocontrol, atención y NFC). Debido a la reducción de las variables introducidas, la variable C, que fue eliminada en el modelo 3 por problemas de colinealidad con otras variables, fue recuperada para este análisis. No obstante, el incremento de varianza explicada por este modelo respecto del modelo 2 (CMG y personalidad) y del modelo 1 (CMG) fue pequeño ($\Delta R^2 = .03$ y $\Delta R^2 = .07$, respectivamente). Los resultados reflejaron que la CMG fue el determinante más robusto de la RC ($\beta = .75$ en el modelo 1, $\beta = .81$ en el modelo 2 y $\beta = .78$ en el modelo 4) y fue la variable que explicó prácticamente la totalidad del desempeño en el TRC.

Los datos también mostraron que, aunque en menor proporción, las subescalas de riesgo ($\beta = .41$), impulsividad ($\beta = -.36$) e intuición ($\beta = .06$) y los factores de personalidad de C y EE ($\beta = .14$ y $\beta = -.06$, respectivamente) también contribuyeron de forma independiente a explicar el éxito en el TRC. Los análisis indicaron que la EX, AP y AM no contribuyeron directamente en la determinación de la RC, pero sí produjeron un efecto indirecto a través de las DP.

La Figura 10 recoge la representación gráfica del modelo 4. El Apéndice L muestra los efectos directos de las variables en la predicción de la RC y el Apéndice M representa gráficamente los efectos indirectos de las variables sobre la RC. Los índices de ajuste del modelo se presentan en la Tabla 58 y fueron calculados mediante el programa LISREL (8.2) (Jöreskog y Sörbom, 1998).

Tabla 56.
Matriz de Correlaciones Corregidas Utilizada en el Análisis de Regresión para Determinar el Desempeño del TRC-13 (Muestra Total)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. WPT	-												
2. Autocontrol	-.060	-											
3. Impulsividad	.187	-.395	-										
4. Atención	.202	.512	-.101	-									
5. Esfuerzo	-.003	.787	-.187	.772	-								
6. Intuición	.139	-.089	.703	.186	.060	-							
7. Riesgo	.118	-.387	.637	-.028	-.184	.463	-						
8. NFC	.422	.218	.220	.529	.418	.468	.305	-					
9. EE	.071	.053	-.111	.144	.007	.003	.174	.058	-				
10. EX	-.053	-.103	.580	-.050	-.022	.356	.458	.065	-.047	-			
11. AP	.232	.230	.105	.120	-.011	.186	.287	.634	-.158	-.183	-		
12. AM	-.173	-.145	-.392	-.111	-.208	-.211	-.019	-.132	-.028	-.267	-.198	-	
13. C	-.164	.519	-.511	.216	.436	-.416	-.762	-.208	-.268	-.572	-.193	-.186	-
14. TRC-13	.746	-.047	.026	.183	.043	.048	.184	.349	.068	-.043	.213	-.031	-.125

Nota. WPT = Wonderlic personnel test; NFC = need for cognition; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a).

Tabla 57.

Análisis de Regresión Múltiple para Determinar el Desempeño del TRC-13 (Muestra Total)

	Modelo 1 (β)	Modelo 2 (β)	Modelo 3 (β)	Modelo 4 (β)
WPT	.75***	.81***	.79***	.78***
EE	-	.17***	-.11***	-.06**
EX	-	.34***	-.04	-
AP	-	.25***	-.07	-
AM	-	.32***	-.07*	-
C	-	.36***	-	.14***
Autocontrol	-	-	-.03	-
Impulsividad	-	-	-.48***	-.36***
Atención	-	-	-.01	-
Esfuerzo	-	-	-	-
Intuición	-	-	.08**	.06*
Riesgo	-	-	.36***	.41***
NFC	-	-	.02	-
<i>F</i>	1,389.14	273.48	175.40	315.57
<i>R</i>	.75***	.77***	.80***	.80***
<i>R</i> ²	.56	.60	.63	.63
<i>R</i> _{CV}	.75	.77	.79	.79
<i>R</i> ² _{CV}	.56	.60	.62	.63

Nota. WPT = Wonderlic personnel test; NFC = need for cognition; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.
 p* < .05; *p* < .01; ****p* < .001

Tabla 58.

Índices de Ajuste del Modelo de Predicción de la RC (Muestra Total)

	χ^2	<i>CFI</i>	<i>RMSEA</i>	<i>RMR</i>
Modelo 4	858.99 (<i>p</i> = .00)	.87	.28	.05

Nota. *CFI* = índice de bondad de ajuste comparativo; *RMSEA* = raíz del residuo cuadrático promedio de aproximación; *RMR* = raíz del residuo cuadrático promedio.

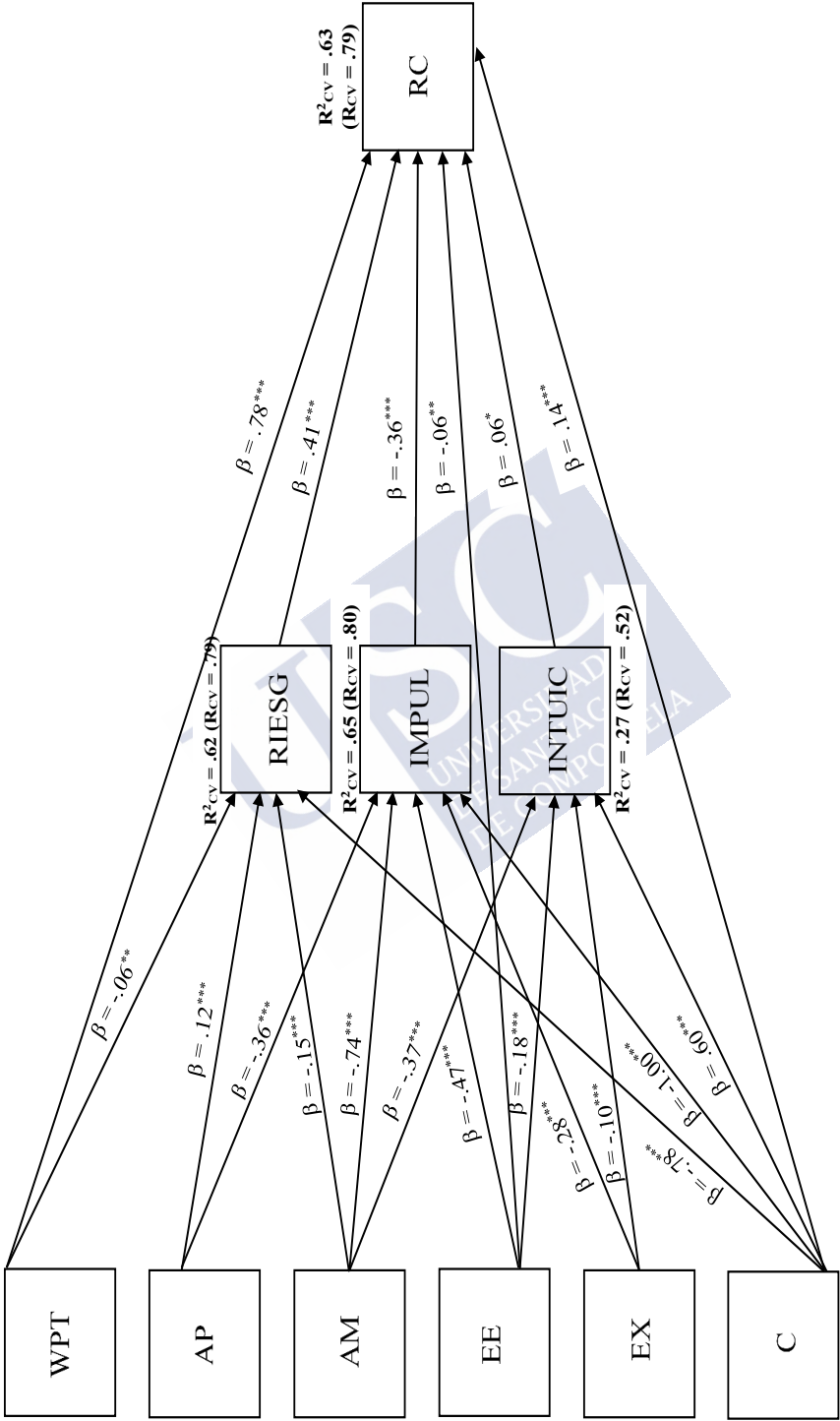


Figura 10. Representación Gráfica del Modelo de Predicción de la RC (Muestra Total)

Otro de los propósitos de este estudio ha sido examinar la predicción múltiple de la RC de forma independiente en la muestra de hombres y de mujeres con el objetivo de comprobar si las diferencias sistemáticas en la RC de los grupos pueden ser debidas a que las variables individuales que determinan este constructo en ambos sexos son diferentes. Por lo tanto, se han llevado a cabo dos nuevos análisis de regresión con las correlaciones corregidas de las muestras de hombres y de mujeres, respectivamente. Para operativizar y facilitar la lectura de este estudio, a continuación, se presentan los resultados más relevantes encontrados al respecto y se relega al Apéndice N toda la explicación detallada a cerca de los pasos llevados a cabo para el análisis de regresión, el resumen pormenorizado de los resultados, así como las tablas con los datos concretos obtenidos y la representación gráfica de los modelos.

En síntesis, los resultados mostraron que tanto variables cognitivas como de personalidad determinan el éxito en el TRC de hombres y mujeres y que, en conjunto, todas ellas explican la misma proporción de varianza en ambos grupos ($R^2_{CV} = .66$ en hombres y $R^2_{CV} = .67$ en mujeres). No obstante, las variables determinantes y la contribución de cada una de ellas parecen diferir en función del sexo. Así, mientras que la CMG se mostró como el determinante más relevante de la RC en mujeres ($\beta = .97$), en el grupo de los hombres fue el cuarto más robusto ($\beta = .62$). En la muestra de hombres, las variables de personalidad (NFC, impulsividad y AP) fueron las variables que más contribuyeron en la explicación del desempeño en el TRC ($\beta = .71$, $\beta = -.68$ y $\beta = -.64$, respectivamente). La DP de riesgo resultó ser un determinante independiente de la RC de las mujeres ($\beta = .28$), pero, sin embargo, en el caso de los hombres no lo fue. Del mismo modo, las DP de autocontrol y atención contribuyeron a explicar varianza de la RC de los hombres ($\beta = -.58$ y $\beta = .16$, respectivamente), pero no de las mujeres. Estos resultados, parecen apoyar la afirmación de Campitelli y Gerrans (2014) de que diferentes variables podrían explicar el desempeño en el TRC en hombres y mujeres.

Validez de Criterio

Reflexibilidad Cognitiva (RC) y Desempeño

Las correlaciones corregidas entre los tres instrumentos de RC y las diversas variables criterio se recogen en la Tabla 59 para las tres muestras examinadas. Los resultados reflejan que la RC fue un predictor válido del desempeño de tarea en el ámbito académico y ocupacional tanto en la muestra total como en el grupo de hombres y de mujeres.

Tabla 59.
Correlaciones Corregidas entre el TRC y las Medidas de Desempeño Académico y Ocupacional

	TRC-13						TRC-10						TRC-3					
	IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)		
	r_{op}	LI	LS	r_{op}	LI	LS	r_{op}	LI	LS	r_{op}	LI	LS	r_{op}	LI	LS	r_{op}	LI	LS
N total																		
<i>D. Acad.</i>																		
CDTE	-.111	-.174	-.048	-.137	-.199	-.074	-.118	-.181	-.055	-.156	-.219	-.094	-.073	-.135	-.010	-.086	-.149	-.024
GPA	.166	.101	.232	.204	.140	.269	.152	.085	.218	.201	.135	.266	.169	.106	.233	.201	.138	.264
Bachiller.	.354	.203	.506	.435	.295	.576	.342	.188	.496	.452	.313	.590	.307	.156	.459	.365	.220	.510
CAU	.370	.244	.496	.454	.339	.570	.361	.233	.489	.476	.363	.590	.306	.179	.433	.363	.242	.485
CDCE	-.076	-.140	-.013	-.095	-.158	-.031	-.085	-.149	-.021	-.112	-.176	-.049	-.044	-.107	.019	-.052	-.115	.011
CDAN	-.037	-.104	.030	-.045	-.112	.022	-.026	-.094	.042	-.035	-.103	.034	-.055	-.120	.011	-.065	-.130	.000
<i>D. Ocup.</i>																		
GCT	.356	.273	.439	.436	.359	.513	.379	.297	.461	.498	.427	.570	.218	.130	.306	.259	.173	.345
Hombres																		
<i>D. Acad.</i>																		
CDTE	-.121	-.235	-.007	-.148	-.261	-.035	-.114	-.229	.001	-.158	-.271	-.044	-.107	-.220	.005	-.123	-.235	-.011
GPA	.095	-.030	.220	.116	-.009	.241	.074	-.053	.200	.102	-.024	.228	.133	.013	.252	.152	.033	.270
Bachiller.	.327	.074	.579	.401	.163	.640	.255	-.004	.515	.357	.113	.601	.470	.227	.712	.539	.318	.760
CAU	.405	.196	.614	.497	.308	.686	.360	.145	.575	.503	.318	.689	.458	.249	.667	.524	.333	.716
CDCE	-.021	-.137	.095	-.026	-.142	.090	-.020	-.136	.096	-.028	-.144	.088	-.024	-.138	.089	-.028	-.141	.086
CDAN	-.083	-.207	.042	-.100	-.224	.023	-.075	-.200	.049	-.105	-.228	.019	-.072	-.194	.049	-.083	-.205	.038

Continúa

Tabla 59.
Continuación

	TRC-13						TRC-10						TRC-3					
	IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)			IC 95% (r_{op})			IC 95% (ρ)		
	r_{op}	LI	LS	ρ	LI	LS	r_{op}	LI	LS	ρ	LI	LS	r_{op}	LI	LS	ρ	LI	LS
<i>D. Ocup.</i>																		
GCT	.354	.205	.504	.433	.294	.571	.349	.198	.499	.483	.351	.614	.316	.158	.474	.362	.210	.514
Mujeres																		
<i>D. Acad.</i>																		
CDTE	-.050	-.128	.027	-.064	-.142	.013	-.066	-.143	.012	-.090	-.167	-.013	-.011	-.088	.066	-.013	-.089	.064
GPA	.226	.149	.303	.289	.215	.363	.212	.134	.290	.290	.216	.365	.211	.130	.293	.254	.181	.328
Bachiller.	.365	.180	.550	.464	.298	.631	.387	.203	.571	.526	.371	.681	.229	.043	.415	.276	.095	.457
CAU	.358	.205	.512	.457	.318	.596	.363	.207	.518	.494	.360	.628	.266	.112	.420	.320	.172	.469
CDCE	-.059	-.137	.018	-.076	-.154	.001	-.073	-.150	.005	-.100	-.177	-.023	-.026	-.103	.051	-.031	-.108	.046
CDAN	-.076	-.157	.005	-.097	-.178	-.016	-.062	-.144	.020	-.086	-.167	-.004	-.096	-.175	-.018	-.115	-.194	-.037
<i>D. Ocup.</i>																		
GCT	.366	.268	.465	.466	.377	.554	.408	.313	.504	.554	.476	.633	.159	.053	.264	.191	.087	.295

Nota. TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); TRC-10 = ítems Salgado (2014a); TRC-3 = ítems Frederick (2005); r_{op} = correlación operativa o correlación corregida por error de medida en el criterio y restricción indirecta en el rango en el predictor; IC 95% (r_{op}) = intervalos de confianza del 95% de la correlación operativa; ρ = correlación verdadera o correlación observada corregida por error de la medida del predictor y criterio y por restricción indirecta en el rango del predictor; IC 95% (ρ) = intervalos de confianza del 95% de la correlación verdadera; LI = límite inferior; LS = límite superior; N total = muestra total; D. Acad. = desempeño académico; D. Ocup. = desempeño ocupacional; CDTE = desempeño de tarea; CDCE = desempeño contextual; CDAN = conductas contraproductivas; GCT = general clerical test; GPA = grade point average; Bachiller = Bachillerato; CAU = calificación de acceso a la universidad.

Respecto al desempeño académico, los datos evidencian que en la muestra total y en el grupo de los hombres, la RC correlacionó de forma débil y negativa con el desempeño de tarea cuando éste fue evaluado con la medida de autoinforme (rango de $\rho = -.156$ a $-.086$ y rango de $\rho = -.158$ a $-.123$, respectivamente) y de forma prácticamente nula en el grupo de mujeres (rango de $\rho = -.090$ a $-.013$). No obstante, la relación entre las variables fue más robusta y positiva cuando el desempeño fue evaluado con medidas objetivas. Así pues, los tamaños del efecto teóricos entre la RC con GPA (rango de $\rho = .201$ a $.204$ en la muestra total; $\rho_{TRC-3} = .152$ en los hombres y rango de $\rho = .254$ a $.290$ en las mujeres), con bachillerato (rango de $\rho_{MT} = .365$ a $.452$, rango de $\rho_H = .357$ a $.539$ y rango de $\rho_M = .276$ a $.526$) y con CAU (rango de $\rho_{MT} = .363$ a $.476$, rango de $\rho_H = .497$ a $.524$ y rango de $\rho_M = .320$ a $.494$) fueron positivas y moderadas, en la mayor parte de los casos. Del mismo modo, los tamaños del efecto teóricos entre la RC y el desempeño ocupacional (tarea de clasificación del GCT) fueron positivos y mayoritariamente moderados en las 3 muestras examinadas (rango de $\rho_{MT} = .259$ y $.498$, rango de $\rho_H = .362$ a $.483$ y rango de $\rho_M = .191$ a $.554$).

En conjunto, estos datos sugieren que, la RC es un predictor válido del desempeño académico y ocupacional, apoyando nuestras Hipótesis 4a y 4b. Además, se examinaron las diferencias entre los coeficientes de correlación del TRC-13 y del TRC-3 y únicamente se encontraron diferencias significativas entre la validez de criterio de ambas medidas para predecir la muestra de trabajo (desempeño ocupacional; $z \geq 3.11$). Los datos sugieren que el TRC-13 es capaz de explicar más varianza de la tarea administrativa que el TRC-3. Asimismo, no se han encontrado diferencias significativas entre la validez de criterio de la RC de hombres y mujeres para predecir el desempeño académico y ocupacional ($z < 1.96$), lo que indica que este constructo es un predictor igualmente válido del desempeño académico y ocupacional en ambos grupos.

Finalmente, en el contexto académico también se ha explorado el desempeño contextual y las conductas contraproductivas. En este sentido, los resultados mostraron que la RC no predice ninguna de esas dos dimensiones del desempeño en ninguna de las 3 muestras. Las correlaciones corregidas entre las tres medidas de RC con el desempeño contextual (rango de $\rho_{MT} = -.112$ a $-.052$, rango de $\rho_H = -.028$ a $-.026$ y rango de $\rho_M = -.100$ a $-.031$) y las conductas contraproductivas (rango de $\rho_{MT} = -.065$ a $-.035$, rango de $\rho_H = -.105$ a $-.083$ y rango de $\rho_M = -.115$ a $-.086$) fueron muy pequeñas o prácticamente nulas. Estos resultados sugieren que la RC podría no ser un predictor del desempeño contextual y de las conductas contraproductivas en el contexto académico.

Validez Añadida

Con el objetivo de examinar si la RC presenta validez añadida sobre otros predictores para explicar el desempeño académico y ocupacional, se han llevado a cabo varios análisis de

regresión con cada medida criterio (notas de bachillerato, CAU, GPA y GCT), empleando el TRC-13 como instrumento de evaluación de la RC. Concretamente, se han calculado 8 modelos de regresión. Investigaciones anteriores han demostrado que la CMG es el mejor predictor del desempeño ocupacional (Bertua, Anderson y Salgado, 2005; Hunter y Hunter, 1984; Salgado et al., 2003; Salgado y Moscoso, 2019) y uno de los más relevantes del desempeño académico (Kuncel, Credé y Thomas, 2005; Postlethwaite, 2011; Richardson et al., 2012; Salgado y Moscoso, 2019), por lo que, en el primer modelo se ha introducido únicamente la medida de CMG (WPT) como variable predictora del desempeño. En el segundo modelo, además de la CMG, se ha introducido el TRC-13 para determinar si la RC añade validez en la predicción de las variables criterio. La segunda variable individual que mejor predice el desempeño, en conjunto con la CMG, es el factor C de personalidad (Salgado y Moscoso, 2008). De esta forma, en el modelo 3 se han introducido el WPT y el factor C de personalidad como predictores del desempeño. Con el objetivo de comprobar la validez que la RC añade sobre CMG y C, se ha testado un cuarto modelo en el que se han introducido las variables WPT, C y TRC-13. En el quinto modelo se han introducido las variables WPT, C y los restantes factores de personalidad (EE, AM, EX y AP) y en el modelo 6, además de estas variables, se ha añadido el TRC-13. La medida de personalidad empleada ha sido el inventario Q15F-Tri, por ser el que mayor tamaño de la muestra presenta. Finalmente, el modelo 7 se compone de las variables WPT, personalidad y DP y el modelo 8, además, añade el TRC-13. La medida de DP de dependencia-independencia de campo no se ha incluido en los análisis de regresión por el limitado tamaño de la muestra. Asimismo, la variable autocontrol también ha sido excluida de los análisis de regresión debido a que presenta problemas de colinealidad.

Para llevar a cabo los análisis de regresión se ha utilizado el paquete estadístico SPSS. A tal efecto, se han empleado las correlaciones corregidas por error en las medidas del predictor y del criterio y, según cada caso, por restricción indirecta en el rango del predictor. De este modo, se garantiza un menor sesgo en la estimación de los parámetros de regresión (Fritz et al., 2016; Hunter y Schmidt, 2015; MacKinnon, 2008). Además, con el objetivo de estimar adecuadamente la validez múltiple, se han calculado los coeficientes de correlación y de determinación múltiple cruzados de la población (R_{cv} y R^2_{cv}). La Tabla 60 recoge la matriz de correlaciones corregidas entre las variables. Esta es la matriz de entrada empleada en los diferentes análisis de regresión.

Tabla 60.

Matriz de Correlaciones Corregidas Utilizada en los Análisis de Regresión para Predecir las Medidas Criterio

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. WPT	-													
2. Autocontrol	-.060	-												
3. Impulsividad	.187	-.395	-											
4. Atención	.202	.512	-.101	-										
5. Esfuerzo	-.003	.787	-.187	.772	-									
6. Intuición	.139	-.089	.703	.186	.060	-								
7. Riesgo	.118	-.387	.637	-.028	-.184	.463	-							
8. NFC	.422	.218	.220	.529	.418	.468	.305	-						
9. EE	.071	.053	-.111	.144	.007	.003	.174	.058	-					
10. EX	-.053	-.103	.580	-.050	-.022	.356	.458	.065	-.047	-				
11. AP	.232	.230	.105	.120	-.011	.186	.287	.634	-.158	-.183	-			
12. AM	-.173	-.145	-.392	-.111	-.208	-.211	-.019	-.132	-.028	-.267	-.198	-		
13. C	-.164	.519	-.511	.216	.436	-.416	-.762	-.208	-.268	-.572	-.193	-.186	-	
14. TRC-13	.746	-.047	.026	.183	.043	.048	.184	.349	.068	-.043	.213	-.031	-.125	-
15. GPA	.262	.276	-.142	.171	.247	-.103	-.196	.194	-.117	-.084	.078	-.121	.268	.204
16. Bachiller.	.557	.096	.056	.039	.091	-.144	.041	.128	-.046	.032	.107	-.190	-.045	.435
17. CAU	.560	.021	.075	.019	.060	-.174	.015	.150	-.064	.012	.104	-.083	-.050	.454
18. GCT	.681	-.010	.139	.145	.043	.087	.021	.158	-.017	-.055	.058	-.160	-.030	.436

Nota. WPT = Wonderlic personnel test; NFC = need for cognition; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; TRC-13 = 3 ítems Frederiek (2005) y 10 ítems Salgado (2014a); GPA = grade point average; Bachiller = Bachillerato; CAU = calificaciones de acceso a la universidad; GCT = general clerical test.

En relación con el desempeño académico, la Tabla 61 muestra los resultados de los análisis de regresión para predecir las notas de bachillerato. Como se puede observar, el TRC-13 no añadió validez sobre otras variables para predecir este criterio académico. El modelo que explicó más varianza de las notas de bachillerato ha sido el modelo 1 ($R^2_{CV} = .31$) compuesto por CMG ($\beta_{WPT} = .56$) y el modelo 7 ($R^2_{CV} = .32$) donde la CMG y las DP de esfuerzo e intuición resultaron ser predictores independientes ($\beta_{WPT} = .56$, $\beta_{Esfuerzo} = .53$ y $\beta_{Intuición} = -.39$). No obstante, cuando a estos modelos se les añadió la variable de RC perdieron, respectivamente, un 3% y un 2% de su capacidad predictiva ($R^2_{CV} = .28$ en el modelo 2 y $R^2_{CV} = .30$ en el modelo 8).

Tabla 61.

Análisis de Regresión para Predecir el Desempeño Académico (Notas de Bachillerato)

	Modelo 1 (β)	Modelo 2 (β)	Modelo 3 (β)	Modelo 4 (β)	Modelo 5 (β)	Modelo 6 (β)	Modelo 7 (β)	Modelo 8 (β)
WPT	.56***	.52***	.57***	.53***	.56***	.50***	.56***	.46**
C	-	-	.05	.05	.02	-.01	-.65	-.77*
EE	-	-	-	-	-.09	-.10	-.26	-.30
AM	-	-	-	-	-.09	-.12	-.32	-.39
EX	-	-	-	-	.03	.01	-.30	-.38
AP	-	-	-	-	-.05	-.07	-.15	-.22
Intuición	-	-	-	-	-	-	-.39**	-.42**
NFC	-	-	-	-	-	-	-.08	-.04
Esfuerzo	-	-	-	-	-	-	.53***	.53***
Impulsividad	-	-	-	-	-	-	.04	.07
Riesgo	-	-	-	-	-	-	-.03	-.07
Atención	-	-	-	-	-	-	-.22	-.21
TRC-13	-	.04	-	.04	-	.08	-	.10
<i>F</i>	59.37	29.58	29.77	19.77	10.45	8.99	9.23	8.54
<i>R</i>	.56***	.56***	.56***	.56***	.58***	.58***	.69***	.69***
<i>R</i> ²	.31	.30	.30	.30	.30	.30	.43	.42
<i>R</i> _{CV}	.56	.53	.53	.52	.49	.48	.57	.55
<i>R</i> ² _{CV}	.31	.28	.28	.27	.24	.23	.32	.30

Nota. WPT = Wonderlic personnel test; C = conciencia; EE = estabilidad emocional; AM = amigabilidad; EX = extroversión; AP = apertura a la experiencia; NFC = need for cognition; TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

p* < .05; *p* < .01; ****p* < .001

La Tabla 62 recoge los resultados de los análisis de regresión empleando como criterio académico las calificaciones de acceso a la universidad (CAU). Los datos nuevamente indicaron que el TRC-13 no añadió validez sobre otras medidas para predecir las CAU. El modelo que explicó más varianza de las CAU ha sido el modelo 7 ($R^2_{CV} = .40$), donde la CMG y las DP de esfuerzo, impulsividad, intuición y atención fueron predictores independientes ($\beta_{WPT} = .67$, $\beta_{Esfuerzo} = .54$, $\beta_{Impulsividad} = .52$, $\beta_{Intuición} = -.49$ y $\beta_{Atención} = -.29$). No obstante, cuando se añadió la

RC al análisis, el modelo continuó explicando la misma proporción de varianza ($R^2_{CV} = .40$), por lo que la RC no contribuyó a determinar las CAU. Sin embargo, en los restantes modelos predictivos de las CAU, cuando se añadió el TRC-13, éstos sufrieron una pérdida de eficiencia predictiva (véase, p. ej., los modelos 3 ($R^2_{CV} = .30$) vs. 4 ($R^2_{CV} = .29$)), lo que sugiere un efecto mediador de la RC. La mediación puede interpretarse como la reducción de la proporción de varianza explicada directamente por las variables independientes en la variable criterio, tras introducir una nueva variable en el modelo (Murgui y Jiménez, 2013). Una revisión detallada del peso de los coeficientes de regresión de esos modelos indica que, cuando se añade el TRC, la β del WPT decrece en todos los modelos y, únicamente, en el modelo 6 también decrecen los pesos de las demás variables independientes. Por lo tanto, el TRC produce un efecto mediador sobre esas variables.

Tabla 62.

Análisis de Regresión para Predecir el Desempeño Académico (CAU)

	Modelo 1 (β)	Modelo 2 (β)	Modelo 3 (β)	Modelo 4 (β)	Modelo 5 (β)	Modelo 6 (β)	Modelo 7 (β)	Modelo 8 (β)
WPT	.56***	.50***	.57***	.51***	.61***	.55***	.67***	.53**
C	-	-	.04	.04	.16	.13	-.05	-.22
EE	-	-	-	-	-.05	-.06	.08	.02
AM	-	-	-	-	.10	.08	.27	.18
EX	-	-	-	-	.17	.14	.12	.00
AP	-	-	-	-	.04	.02	.29	.20
Intuición	-	-	-	-	-	-	-.49***	.53***
NFC	-	-	-	-	-	-	-.20	-.15
Impulsividad	-	-	-	-	-	-	.52***	.56***
Esfuerzo	-	-	-	-	-	-	.54***	.54***
Atención	-	-	-	-	-	-	-.29**	-.26**
Riesgo	-	-	-	-	-	-	-.20	-.27*
TRC-13	-	.08	-	.08	-	.07	-	.13
<i>F</i>	87.26	44.00	43.77	29.43	15.45	13.27	14.96	14.04
<i>R</i>	.56***	.56***	.56***	.56***	.58***	.58***	.71***	.71***
<i>R</i> ²	.31	.31	.31	.31	.31	.31	.47	.47
<i>R</i> _{CV}	.55	.55	.55	.54	.52	.51	.63	.63
<i>R</i> ² _{CV}	.31	.30	.30	.29	.27	.26	.40	.40

Nota. WPT = Wonderlic personnel test; C = conciencia; EE = estabilidad emocional; AM = amigabilidad; EX = extroversión; AP = apertura a la experiencia; NFC = need for cognition; TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

p* < .05; *p* < .01; ****p* < .001

La Tabla 63 sintetiza los resultados de la variable GPA. Los datos mostraron que la RC añadió validez sobre las medidas cognitivas y de personalidad para predecir el GPA. Conjuntamente CMG, los factores de personalidad y las DP explicaron el 24% de su varianza

($R_{CV} = .49$). No obstante, cuando a este modelo se le añadió la variable de RC (modelo 8), los predictores, de forma acumulada, explicaron el 25% de la varianza del GPA ($\Delta R^2_{CV} = .01$). Por lo tanto, la RC contribuyó en la predicción múltiple del GPA.

Tabla 63.

Análisis de Regresión para Predecir el Desempeño Académico (GPA)

	Modelo 1 (β)	Modelo 2 (β)	Modelo 3 (β)	Modelo 4 (β)	Modelo 5 (β)	Modelo 6 (β)	Modelo 7 (β)	Modelo 8 (β)
WPT	.26***	.25***	.31***	.30***	.39***	.46***	.52***	.74***
C	-	-	.32***	.32***	.83***	.86***	1.19***	1.45***
EE	-	-	-	-	.17***	.18***	.32***	.40***
AM	-	-	-	-	.32***	.35***	.50***	.64***
EX	-	-	-	-	.57***	.60***	.81***	1.00***
AP	-	-	-	-	.34***	.36***	.54***	.67***
Esfuerzo	-	-	-	-	-	-	.09	.08
Riesgo	-	-	-	-	-	-	.10	.21**
NFC	-	-	-	-	-	-	-.10	-.19*
Atención	-	-	-	-	-	-	.24***	-.28***
Impulsividad	-	-	-	-	-	-	-.06	-.12
Intuición	-	-	-	-	-	-	.12*	.18**
TRC-13	-	.02	-	.02	-	-.09	-	-.21***
<i>F</i>	65.45	32.77	89.56	59.73	47.75	41.56	26.44	26.04
<i>R</i>	.26***	.26***	.41***	.41***	.50***	.50***	.52***	.53***
<i>R</i> ²	.07	.07	.17	.17	.24	.24	.26	.27
<i>R</i> _{CV}	.26	.26	.41	.41	.48	.48	.49	.50
<i>R</i> ² _{CV}	.07	.07	.18	.17	.23	.23	.24	.25

Nota. WPT = Wonderlic personnel test; C = conciencia; EE = estabilidad emocional; AM = amigabilidad; EX = extroversión; AP = apertura; NFC = need for cognition; TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

* $p < .05$; ** $p < .01$; *** $p < .001$

Finalmente, se muestran los resultados de la predicción múltiple del desempeño ocupacional. Como se puede observar en la Tabla 64, la RC añadió validez sobre CMG (modelo 2) y sobre CMG, NFC y esfuerzo (modelo 8) para explicar la muestra de trabajo administrativa (GCT; $\Delta R^2_{CV} = .01$, en ambos casos). No obstante, en ambos modelos los pesos de los coeficientes de regresión del TRC-13 fueron negativos a pesar de la correlación positiva que existe entre ambos constructos ($\beta = -.16$, en ambos casos), lo que sugiere que el TRC funciona como una variable supresora en estos modelos. Las variables supresoras son aquellas que, al introducirlas en el modelo, incrementan el coeficiente de determinación de las otras variables en el pronóstico del criterio (Cohen y Cohen, 1983; Conger, 1974; Horst, 1941; McKinnon, Lockwood y Williams, 2007). Una revisión detallada de los coeficientes de regresión indica que el peso de las β aumentó

al incluir el TRC-13 ($\Delta\beta_{\text{WPT}} = .17$, $\Delta\beta_{\text{NFC}} = .06$ del modelo 8 respecto del modelo 7; $\Delta\beta_{\text{WPT}} = .12$ del modelo 2 respecto del modelo 1).

Tabla 64.

Análisis de Regresión para Predecir el Desempeño Ocupacional (GCT)

	Modelo 1 (β)	Modelo 2 (β)	Modelo 3 (β)	Modelo 4 (β)	Modelo 5 (β)	Modelo 6 (β)	Modelo 7 (β)	Modelo 8 (β)
WPT	.68***	.80***	.70***	.82***	.68***	.77***	.75***	.92**
C			.08*	.08*	-.15*	-.11	-.22	-.02
EE					-.15**	-.13**	-.13	-.06
AM					-.17**	-.13*	-.10	.00
EX					-.20**	-.16*	-.18	-.04
AP					-.22***	-.19***	-.06	.05
NFC							-.29**	-.35***
Atención							.03	.00
Impulsiv.							.03	-.02
Intuición							.06	.11
Esfuerzo							.21**	.21**
Riesgo							-.03	.05
TRC-13		-.16**		-.16**		-.11*		-.16*
<i>F</i>	407.34	213.01	208.91	145.61	76.09	66.42	40.94	38.95
<i>R</i>	.68***	.69***	.69***	.69***	.71***	.71***	.72***	.72***
<i>R</i> ²	.46	.47	.47	.48	.49	.49	.50	.51
<i>R</i> _{CV}	.68	.68	.68	.69	.69	.69	.69	.70
<i>R</i> ² _{CV}	.46	.47	.47	.47	.48	.48	.47	.48

Nota. WPT = Wonderlic personnel test; C = conciencia; EE = estabilidad emocional; AM = amigabilidad; EX = extroversión; AP = apertura a la experiencia; NFC = need for cognition; Impulsiv. = impulsividad; TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

p* < .05; *p* < .01; ****p* < .001

En resumen, los resultados mostraron que la RC añadió validez sobre otras variables individuales en la predicción del desempeño ocupacional (muestra de trabajo) y del desempeño académico (GPA). En el caso del desempeño académico, la RC únicamente contribuyó a la predicción del GPA, pero no de las otras variables criterio (bachillerato y CAU). Aunque los análisis de correlaciones bivariadas han mostrado que el TRC-13, por sí solo, es un predictor válido de las notas de bachillerato ($\rho_{\text{TRC-13}} = .435$), de las CAU ($\rho_{\text{TRC-13}} = .454$), el análisis de predicción múltiple sugiere que, cuando se controla por los efectos de otras variables disposicionales (variables cognitivas y de personalidad), la RC no contribuye de forma independiente a explicar las notas de bachillerato ni las CAU.

4. DISCUSIÓN

El principal objetivo de este estudio ha sido conocer la validez de constructo y criterio del TRC. Para ello, en primer lugar, se ha examinado la validez convergente y discriminante y se ha tratado de determinar las variables individuales que de forma conjunta explican el desempeño en el TRC. En segundo lugar, se ha examinado la validez de criterio y se ha explorado si la RC añade validez sobre otras variables disposicionales para predecir criterios académicos y ocupacionales.

Los resultados han mostrado que la RC es un constructo, principalmente, de capacidad cognitiva y, en menor medida, de personalidad (factores de personalidad y DP). De acuerdo con los resultados, la estimación teórica de la relación entre RC y CMG ha oscilado entre .54 y .74, dependiendo de la medida de CMG empleada. La magnitud de la relación teórica entre la RC y las variables de personalidad ha sido menor. Únicamente las dimensiones de EE, AP y C han demostrado asociarse de forma débil con RC. No obstante, la extraversión y la amigabilidad producen un efecto indirecto a través de las DP que determinan el constructo.

Simultáneamente, los resultados han reflejado una relación entre la RC y la disposición hacia un pensamiento (1) de compromiso y disfrute con la cognición, (2) de atención y concentración en la realización de tareas, (3) de independencia de campo y (4) de no aversión al riesgo. Estos resultados apoyan las conclusiones de investigaciones previas. Por ejemplo, Frederick y otros autores (2005; Haran et al., 2013; Primi et al., 2017; Thomson y Oppenheimer, 2016; Toplak et al., 2014) encontraron evidencias de que la disposición a comprometerse y disfrutar con la realización de tareas que implican esfuerzo cognitivo determina la RC. Del mismo modo, prestar atención en la realización de la tarea y ser capaces de mantenerse abstraído de cualquier estímulo distractor durante la realización del TRC podría permitir a los sujetos interpretar mejor el significado de los ítems y detectar más fácilmente las respuestas capciosas. De acuerdo con Macchi y Bagassi (2012), las diferencias individuales en la RC podrían ser debidas a un tipo muy concreto de flexibilidad, esto es, a la competencia pragmática o pragmalingüística, ya que la tarea se interpreta en el contexto de presentación de la información del enunciado y no se abstrae o se descontextualiza de él. En este sentido, la disposición hacia un pensamiento independiente de campo desempeñaría un rol importante, al permitir “romper” el campo organizado del enunciado y abstraer de él la parte capciosa que activa el procesamiento de tipo 2.

La segunda contribución de este estudio ha sido mostrar que las variables que determinan la RC de hombres y mujeres no son exactamente las mismas. Concretamente, se han encontrado diferencias en la capacidad predictiva, considerada individualmente, de (1) la DP de riesgo y (2) los factores de personalidad de amigabilidad y conciencia. Los resultados indicaron que la disposición a asumir riesgos puede explicar en mayor proporción la varianza de la RC de hombres

que de mujeres y que los factores de amigabilidad y conciencia únicamente determinan el desempeño en el TRC de los hombres.

Sin embargo, los resultados de la predicción múltiple de la RC de hombres y de mujeres muestran diferencias en la contribución, no solo de las variables de personalidad, sino también de las variables cognitivas. Por ejemplo, mientras que el factor de personalidad EE y las DP de autocontrol y atención contribuyen de forma independiente a la predicción múltiple de la RC de los hombres, no añaden validez en el caso de las mujeres (salvo EE que produce un efecto indirecto). Del mismo modo, mientras que el factor de personalidad EX y las DP de esfuerzo y riesgo contribuyen de forma independiente a la predicción múltiple de la RC de las mujeres, no añaden validez en el caso de los hombres (salvo EX que produce un efecto indirecto). Igualmente, la CMG se presenta como el determinante más robusto de la RC de las mujeres, mientras que es el cuarto más relevante en la muestra de hombres. El resto de las variables disposicionales examinadas contribuyen de forma diferente a la explicación del constructo en ambos grupos. A pesar de ello, la proporción de varianza que las variables predictoras en conjunto explican de la RC de hombres y mujeres es prácticamente la misma.

El segundo objetivo de esta investigación ha sido examinar la validez de criterio de la RC y explorar la validez que añade sobre otros predictores del desempeño ocupacional y académico. En este sentido, la tercera contribución ha sido mostrar que la RC es un predictor relevante del desempeño académico y ocupacional. De acuerdo con los resultados, la estimación teórica de la relación entre la RC y el desempeño académico ha sido .204 para GPA, .435 para las notas de bachillerato y .454 para las CAU. Asimismo, la estimación teórica de la relación entre la RC y el desempeño ocupacional ha sido .436. Estos resultados tienen implicaciones para la práctica profesional de la selección de personal en las empresas y los procesos de admisión de alumnos a estudios no obligatorios, ya que los resultados sugieren que el TRC es un instrumento válido para la toma de decisiones en ambos contextos.

En relación con la validez de criterio, este estudio también ha examinado la validez de la RC para predecir el desempeño contextual y las conductas contraproductivas. En este sentido, la cuarta contribución ha sido mostrar que la RC predice el desempeño de tarea, pero no el desempeño contextual ni las conductas contraproductivas en el contexto académico. A nuestro juicio, ningún otro estudio había examinado previamente esta cuestión, por lo que ésta sería una contribución única de esta investigación. En cualquier caso, nuevos estudios deberían tratar de replicar estos resultados.

Por último, la quinta contribución de este trabajo ha sido mostrar indicios de que la RC añade validez sobre CMG, los factores de personalidad y las DP para predecir el desempeño académico (medido con GPA) y el desempeño ocupacional. Aunque el incremento de la validez predictiva de la RC ha sido pequeño, estos resultados tienen importantes implicaciones teóricas: (1) añadiendo la RC al modelo, podemos explicar mayor proporción de varianza del desempeño

y (2) los datos indican que, aunque RC es un constructo principalmente cognitivo, que se asocia de forma moderada con CMG, ambos constructos no son equivalentes. De este modo, los resultados apoyan nuevamente la hipótesis de que la RC no es, únicamente, una variable de capacidad cognitiva, sino que, además, captura varianza de personalidad y disposición a pensar analíticamente (Liberali et al., 2012; Moritz et al., 2014; Pennycook y Ross, 2016; Sirota y Juanchich, 2011; Toplak et al., 2011, 2014).

En resumen, esta investigación ha permitido profundizar en el conocimiento y definición del constructo de la RC. Los datos indican que la RC es un concepto cognitivo que también viene determinado por aspectos de la personalidad. Además, se ha demostrado que, en el caso de hombres y mujeres, aunque las variables disposicionales contribuyen de forma diferente en la explicación de la RC, predicen la misma proporción de varianza en ambos sexos. Asimismo, se ha podido demostrar que la RC es un predictor relevante del desempeño y que añade validez sobre otros predictores importantes.

Para finalizar, deben señalarse las limitaciones de este estudio. En primer lugar, la muestra de hombres es sustancialmente limitada en comparación con la muestra de mujeres, lo que obliga a tomar con mayor precaución los resultados obtenidos para este grupo. Además, para algunas medidas (como el RLS, Factor G o GCT) la proporción de sujetos que completaron esos tests ha sido sustancialmente inferior, lo que ha impedido tenerlos en cuenta para efectuar algunos cálculos (como los análisis de regresión). Próximas investigaciones deberían ir encaminadas a seguir examinando la relación entre la RC y estas variables. Por otro lado, el criterio de desempeño ocupacional fue evaluado con una medida *proxy* del desempeño. Nuevos estudios deberían explorar la validez de la RC para predecir otras medidas de desempeño ocupacional. Asimismo, otras investigaciones deberían ir encaminadas a examinar la validez de la RC para predecir el desempeño contextual y las conductas contraproductivas tanto en el ámbito académico como ocupacional. Finalmente, debido al limitado tamaño de la muestra de hombres, no se ha podido examinar la validez añadida de la RC de hombres y mujeres para predecir las medidas de desempeño académico y ocupacional. Esta es una cuestión que queda pendiente para futuras investigaciones.



ESTUDIO 4.

Meta-Análisis de la Relación de la RC, Capacidad Mental General y Desempeño Académico

1. INTRODUCCIÓN

Con el objetivo de mostrar nuevas evidencias de validez del TRC, se han llevado a cabo diversos meta-análisis. Concretamente, se han hechos tres integraciones cuantitativas para explorar las siguientes cuestiones teóricas: (1) conocer la correlación teórica entre la capacidad cognitiva y la RC, (2) examinar la validez operativa y verdadera de la RC para predecir el desempeño académico y (3) determinar el tamaño del efecto teórico de las diferencias entre hombres y mujeres en RC.

Capacidad Cognitiva y RC

Una de las principales conclusiones del tercer estudio de esta tesis ha sido que el constructo de RC es una variable, principalmente, de capacidad cognitiva. Por este motivo, el primer objetivo de esta investigación ha sido examinar la relación teórica entre la capacidad cognitiva y la RC.

Estudios previos han mostrado una asociación positiva entre múltiples medidas de capacidad cognitiva y RC (Finucane y Gullion, 2010; Frederick, 2005; Primi et al., 2015; Sinayev y Peters, 2015; Thomson y Oppenheimer, 2016). Asimismo, se ha encontrado que esta relación es mayor con medidas de capacidad numérica (r en torno a .40 - .50) y de CMG (r en torno a .30 - .50) que con otras medidas cognitivas. No obstante, hasta el momento, no se ha llevado a cabo ninguna integración cuantitativa de los resultados. El primer meta-análisis realizado en este estudio examina la relación de la CMG, de capacidades cognitivas específicas (como la capacidad verbal, numérica y mecánico-espacial) y la capacidad de la memoria operativa con la RC. Las hipótesis formuladas al respecto son:

Hipótesis 1a: *La CMG se relaciona con la RC de forma positiva y generaliza su validez.*

Hipótesis 1b: *Las capacidades cognitivas específicas se relacionan con la RC de forma positiva y generalizan su validez.*

Hipótesis 1c: *La capacidad de la memoria operativa se relaciona con la RC de forma positiva y generaliza su validez.*

RC y Desempeño Académico

Las evidencias de validez de la RC para predecir el desempeño académico han sido examinadas por algunos investigadores. Resultados previos han mostrado una relación positiva entre la RC y diversas medidas criterio, tales como, GPA (*Grade Point Average*; Agnew y Harrison, 2017; Borghans y Golsteyn, 2014; Insler et al., 2015), el logro matemático (Gómez-Chacón et al., 2014; Lindeman y Svedholm-Häkkinen, 2016; Primi et al., 2017), las calificaciones en física (Lindeman y Svedholm-Häkkinen, 2016) y las calificaciones del examen A de alemán (Lohse, 2016). En base a estas evidencias, la hipótesis propuesta es:

Hipótesis 2: *La RC predice el desempeño académico de forma positiva y generaliza su validez.*

Diferencias entre Hombres y Mujeres en RC

Los resultados de meta-análisis anteriores han sugerido que los hombres obtienen sistemáticamente mejores resultados que las mujeres en el TRC (Brañas-Garza et al., 2015; Cueva et al., 2015; Primi et al., 2018). No obstante, estas integraciones cuentan con una serie de limitaciones que hacen necesario la realización de un nuevo meta-análisis sobre las diferencias en función del sexo. Así, por ejemplo, los estudios de Cueva et al. (2015) y Primi y colaboradores (2018) no tenían por objetivo desarrollar un meta-análisis sobre las diferencias entre hombres y mujeres y, por tanto, no se basan en una exhaustiva revisión de la literatura. Estas revisiones han integrado un conjunto de 8 ($N = 1,180$) y 13 muestras ($N = 2,536$), respectivamente. En los dos casos se trata de estudios desarrollados por el grupo de investigación al que pertenecen los autores, donde el número de casos integrados es muy limitado. En el caso de las investigaciones de Cueva et al. (2015) y de Brañas-Garza y colaboradores (2015), los resultados no informan del tamaño del efecto de esas diferencias, ni aportan datos suficientes para su cálculo. Simplemente se limitan a transcribir la existencia de diferencias en RC en función del sexo y a indicar las medias alcanzadas por ambos grupos. Finalmente, en ninguno de los 3 estudios se han llevado a cabo correcciones por artefactos, de modo que la estimación meta-analítica es una mera aproximación al tamaño del efecto verdadero. El mejor estimador del tamaño del efecto verdadero es aquel en el que las correlaciones observadas han sido corregidas por todas las fuentes de error posibles.

Por lo tanto, el objetivo de este tercer meta-análisis es ampliar las integraciones anteriores y tratar de hacer una estimación más precisa de las diferencias verdaderas entre hombres y mujeres en RC. De acuerdo con los resultados encontrados en la literatura científica, la hipótesis planteada al respecto es:

Hipótesis 3: *Los hombres obtienen puntuaciones más altas que las mujeres en RC.*

Como se ha venido explicando a lo largo de esta tesis, el TRC-3 desarrollado por Frederick (2005) es una medida de RC metodológicamente dudosa, en el sentido de que una serie de limitaciones psicométricas afectan al test y ponen en tela de juicio su validez como instrumento de evaluación de la RC. Consecuentemente, se han desarrollado nuevos test con el objetivo de superar dichas limitaciones. Este estudio trata de examinar las evidencias de validez de la medida de RC diferenciando entre el TRC original y los nuevos instrumentos. Por este motivo, las tres integraciones cuantitativas se han repetido controlando el tipo de medida de RC (TRC original vs. nuevos TRC).

2. PROCEDIMIENTO

Revisión de la Literatura

Para obtener las muestras de los estudios originales y ser incluidos en el meta-análisis se usaron diferentes estrategias de búsqueda. En primer lugar, se realizó una búsqueda en la base de datos ERIC y en los meta-buscadores de Google y Google Scholar. Para ello se introdujeron los términos “*cognitive reflection*” y “*Cognitive Reflection Test*”. Seguidamente, se hizo una búsqueda artículo por artículo en las siguientes revistas científicas: *Applied Cognitive Psychology*, *Cognition*, *Cognitive Science*, *Frontiers in Psychology*, *Journal in Applied Research*, *Journal of Behavioral Decision Making*, *Journal of Economic Behavior and Operation*, *Journal of Experimental Psychology: General*, *Journal of Operations Management*, *Judgment and Decision Making*, *Memory and Cognition*, *Mind and Society*, *Production and Operations Management*, *The Journal of Economic Perspectives*, *The Journal of Socio-Economics* (desde 1995 hasta 2014) y *Journal of Behavioral and Experimental Economics* (desde 2014) y *Thinking and Reasoning*. Teniendo en cuenta que el TRC se publicó por primera vez en 2005, en ambos casos, la búsqueda se ciñó al periodo comprendido desde septiembre de 2005 hasta septiembre de 2018. En tercer lugar, se examinó la sección de las referencias bibliográficas de los diferentes artículos para identificar nuevos estudios. Finalmente, se contactó personalmente con diversos autores con el objetivo de obtener información suplementaria no facilitada en el artículo de origen.

Criterios de Inclusión

Una vez finalizada la fase anterior se procedió al examen de cada uno de los estudios primarios obtenidos con el objetivo de determinar su inclusión y de codificar cada uno de los datos.

Teniendo en cuenta que el objetivo del estudio era conocer el tamaño del efecto verdadero de la relación de la RC con CMG y desempeño académico, así como las diferencias en función del sexo en la RC, para ser incluidos los estudios en el meta-análisis fue necesario (1) que los estudios primarios proporcionasen un indicador de cualquiera de esas relaciones u (2) otro dato que permitiese calcular el tamaño del efecto (medias y *SD*, estadísticos descriptivos, etc.).

Asimismo, se integraron estudios donde el constructo de RC era medido con el TRC original (Frederick, 2005; Kahneman y Frederick, 2002) o con nuevas versiones desarrolladas de este test. Por versiones desarrolladas se entienden: (1) los test de RC que, aun incluyendo 3 o menos ítems originales de Frederick, incluyesen nuevos ítems y (2) TRC de más de 3 ítems compuestos en su totalidad por nuevos ítems. A estos efectos, los TRC compuestos por tres ítems paralelos a los originales no fueron considerados como nuevas versiones del TRC, ya que la diferencia entre estos ítems y los originales radica en la sustitución del sustantivo del problema (por ejemplo, bate y pelota por televisión y CD). La Tabla 65 sintetiza los estudios integrados en los meta-análisis que han empleado nuevas versiones del TRC.

Tabla 65.

Estudios Integrados que han Empleado Nuevas Versiones del TRC

Referencia	N.º de ítems del TRC
Finucane y Gullion (2010)	6
Gómez-Chacón, García-Madruga, Vila, Elosúa y Rodríguez (2014)	5
Grossman, van der Weele y Andrijevich (2014)	4
Toplak, West y Stanovich (2014)	7
Corgnet, Espín y Hernán-González (2015)	7
Pennycook, Cheyne, Barr, Koehler y Fugelsang (2015)	7
Primi, Morsanyi, Chiesi, Donati y Hamilton (2015)	6
Royzman, Landy y Leeman (2015)	6
Sinayev y Peters (2015)	5
Corgnet, Espín y Hernán-González (2015)	7
Jackson, Kleitman, Howie y Stankov (2016)	7
Thomson y Oppenheimer (2016)	7
Morsanyi, McCormack y O'Mahony (2017)	6

Continúa

Tabla 65.

Continuación

Referencia	N.º de ítems del TRC
Primi, Morsanyi, Donati, Galli y Chiesi (2017)	6
Szaszi, Szollosi, Palfi y Aczel (2017)	6
Toplak, West y Stanovich (2017)	11
Yilmaz y Saribay (2017)	4
Primi, Donati, Chiesi y Morsanyi (2018)	6
Estudio 3 de esta tesis	13

En el caso de las variables de capacidad cognitiva, se incluyeron los estudios que relacionasen la RC con medidas de inteligencia general, de capacidad cognitiva específicas y otras medidas cognitivas. De forma que se admitieron medidas de CMG, de capacidad verbal, de capacidad numérica, de capacidad mecánico-espacial y de capacidad de memoria operativa. Cada uno de los instrumentos de inteligencia fueron clasificados en las categorías anteriores. En caso de duda, se acudió al manual del instrumento y se examinaron las evidencias de validez convergente y discriminante. Asimismo, las pruebas cognitivas estandarizadas de admisión académica (p. ej. SAT y ACT) fueron incluidas en la categoría de medidas de CMG. Estudios previos han indicado que la correlación entre las puntuaciones informadas por los participantes y las verificadas u oficiales oscila entre .80 y .90 (Kuncel et., 2005; West et al. 2012), por lo que ambos casos fueron aceptados para su integración en el meta-análisis.

Durante el proceso de recogida de datos se detectaron dos estudios que proporcionaban una correlación entre la RC y la velocidad perceptual. Sin embargo, debido a la falta de investigación primaria, no se pudieron incluir en los análisis. La Tabla 66 presenta una lista de los instrumentos de capacidad cognitiva utilizados en los estudios primarios.

Tabla 66.

Instrumentos de Inteligencia Utilizados en los Estudios Primarios

Instrumentos de Inteligencia
<i>Capacidad Mental General:</i>
Wonderlic Personnel Test (Dodrill, 1983 y otras versiones)
Swedish Psychometric Company Assessio (Sjöberg et al., 2006)
Raven's Advance Progressive Matrices (Raven et al., 1998 y otras versiones)
Location Test (Ekstrom et al., 1976)
Wechsler Abbreviated Scales of Intelligence (Wechsler, 1999)
Employee Aptitude Survey (Ruch et al., 2001)

Continúa

Tabla 66.

Continuación

Instrumentos de Inteligencia
<p><i>Capacidad Mental General:</i></p> <p>Shipley Institute of Living Scale (Shipley (1986)</p> <p>Number Series Test (Arendasy y Sommer, 2012)</p> <p>Number Line Task (Siegler y Opfer, 2003)</p> <p>Ordering abilities (Morsanyi et al., 2017)</p> <p>Pruebas estandarizadas de acceso a la universidad (por ejemplo, SAT o ACT)</p> <p>Prueba de acceso al servicio militar en Israel (<i>Kaba</i>)</p> <p><i>Capacidad Numérica:</i></p> <p>Numeracy Scale de 10 ítems (Lipkus et al., 2001)</p> <p>Numeracy Scale de 3 ítems (Schwartz et al., 1997)</p> <p>Numeracy Scale de 11 ítems (Lipkus et al., 2001; Peters y Levin, 2008)</p> <p>Berlin Numeracy Test (Cokely et al., 2012)</p> <p>Numeracy Scale de 15 ítems (Peters et al., 2007)</p> <p>Numerical Ability Test (Differential Aptitudes Test de Bennett et al., 1989)</p> <p>Rach Based Numeracy Ability (Weller et al. 2013)</p> <p>The Math Fluency Subtest of Woodcock-Johnson III Test of Achievement (Woodcock et al., 2001)</p> <p><i>Capacidad Verbal:</i></p> <p>Recognition Vocabulary (V-1) (Ekstrom et al., 1976).</p> <p>Synonyms and Antonyms (Salthouse, 1993)</p> <p>Shipley Vocabulary Test (Shipley, 1986)</p> <p>Vocabulary (Lindenberger et al., 1993)</p> <p>Mill Hill Vocabulary Scale (Raven et al., 1998)</p> <p>WordSum Task (Huang y Hauster, 1998)</p> <p>Vocabulary Test (Knežević y Opacić, 2011)</p> <p>Analogies Test (Wolf et al., 1992)</p> <p>Synonyms and Antonyms (Wolf et al., 1992)</p> <p>Esoteric Analogies Test (Stankov, 1997)</p> <p><i>Capacidad Mecánico-Espacial:</i></p> <p>Mental Rotation Test (Shepard y Metzler, 1971)</p> <p>Mental Rotation Test (Vanderberg y Kuse, 1978)</p> <p>Mechanical Comprehension Test (Bennett, 1969)</p> <p>3D Space Test (Wolf et al., 1992)</p> <p>Physical Prediction Questionnaire (Lawson et al., 2004)</p> <p><i>Memoria Operativa:</i></p> <p>Backward Digit Span (Wechsler Adult Intelligence Scale III, Wechsler 1987)</p> <p>Paired Associates (Salthouse y Babcock, 1991)</p> <p>Paced Auditory Serial Addition Test (Gronwall, 1977)</p> <p>Working Memory Capacity Span (Turner y Engle, 1989)</p>

Continúa

Tabla 66.

*Continuación**Memoria Operativa:*

Working Memory Capacity Span (Unsworth et al. 2005)

Semantic Anaphora Working Memory Test (Carriedo et al., 2011)

Swaps Test (Stankov, 2000)

Gospan Working Memory Capacity (De Neys et al., 2002)

Los estudios integrados en el meta-análisis de desempeño académico incluyeron medidas tales como el GPA y las calificaciones de exámenes (matemáticas, física, estadística, examen de nivel A de alemán). En estos casos también se aceptaron las medidas auto informadas y las oficiales. Del mismo modo, se incluyeron dentro de esta categoría las pruebas de conocimientos pre-requisito para iniciar un curso.

En relación con las diferencias entre hombres y mujeres en el TRC, se eliminaron algunos estudios que, aunque indicaban que existían diferencias, no reportaban ningún dato que permitiese estimar el tamaño del efecto o los datos facilitados no eran suficientes para hacerlo (p. ej. Ibanez, Riener y Rai, 2013; Corgnet, Espín, Hernán-Gonzalez et al., 2015). Asimismo, se excluyeron aquellos estudios donde los resultados sobre las diferencias entre hombres y mujeres en la RC no eran fácilmente interpretables o no quedaba claro el estadístico empleado para calcular esas diferencias (p. ej. Nieuwenstein y Van Rijn, 2012). Del mismo modo, se excluyeron los estudios en los que se impuso un límite temporal para completar el TRC, ya que esto puede añadir una fuente de error adicional y no común a los resultados de la integración (p. ej. Ring et al., 2016). El objetivo del meta-análisis es estimar cuánta varianza observada es debida a los errores artificiales. Si no se controlan las fuentes de error no comunes entre los estudios singulares, se estará inflando la variabilidad real.

Finalmente, en el meta-análisis sobre la relación entre el TRC y CMG, con una muestra total acumulada de 12,490 sujetos, se han integrado 45 tamaños del efecto. El meta-análisis sobre la relación entre el TRC y desempeño académico se ha ejecutado incluyendo 11 tamaños del efecto y acumulando una muestra total de 3,063 individuos. Por último, en el meta-análisis de diferencias en la RC en función del sexo se ha acumulado una muestra total de 30,529 sujetos y se han integrado 66 tamaños del efecto (véanse Apéndices Ñ-R).

Codificación de Estudios

Cada estudio fue codificado de forma individual en una base de datos. Siempre que el estudio facilitase los datos, la información recopilada fue la siguiente:

1. Referencia del artículo: autores y año de publicación.

2. Información sobre la muestra: tamaño de la muestra (inicial y final, para el total de la muestra y por sexos) y características sociodemográficas como la edad (media, desviación típica y rango), tipo de muestra (estudiantes, trabajadores o población en general) y nacionalidad de la muestra.
3. Información sobre la variable de RC: instrumento utilizado (original o nuevo), número de ítems del test, fiabilidad del test, media y desviación típica de esa muestra en el test y particularidades de la administración del test (por ejemplo, imposición de un límite temporal).
4. Información sobre la variable de capacidad cognitiva: nombre del instrumento utilizado y referencia (autor y año), tipo de inteligencia evaluada (inteligencia general o específica), fiabilidad del instrumento para esa muestra y media y desviación típica de la muestra en esa medida.
5. Información sobre la variable de desempeño académico: nombre de la medida, breve explicación de la misma, fiabilidad de la medida para la muestra y desviación típica de la muestra.
6. Información sobre el tamaño del efecto: se registró el tamaño del efecto, expresado principalmente en r de Pearson, u otros datos que permitiesen calcularlo. Por ejemplo, para el meta-análisis de diferencias se registraron datos como la media y desviación típica de cada grupo, la t de *Student*, el nivel de significatividad de la t de *Student*, d de Cohen, etc.).

Reglas de Decisión

Una vez codificados los estudios singulares y antes de llevar a cabo las integraciones cuantitativas, se han tenido que establecer una serie de reglas ante determinadas particularidades de los estudios primarios. A continuación, se explican las reglas de decisión adoptadas y el procedimiento general seguido para el cálculo de los meta-análisis.

1. Para examinar la relación entre capacidad cognitiva y RC se llevaron a cabo diferentes meta-análisis de correlaciones en función de la medida de inteligencia. De modo que las integraciones llevadas a cabo fueron: (1) con las medidas de capacidad verbal, (2) con las medidas de capacidad numérica, (3) con los instrumentos de capacidad mecánico-espacial, (4) con los test de capacidad de memoria operativa y (5) con los instrumentos de CMG. En este último meta-análisis, se integraron todas las correlaciones entre RC y las medidas de capacidad cognitiva general y las correlaciones entre RC y las medidas de capacidades específicas (entendidas como medidas imperfectas del constructo), tales como

capacidad verbal, numérica y mecánico espacial. En este meta-análisis no se integraron las relaciones entre RC y las medidas de memoria operativa ni velocidad perceptual por no considerarse estas medidas de CMG.

2. Para llevar a cabo el meta-análisis de correlaciones entre el TRC y CMG se integró un único coeficiente de validez por cada muestra. No obstante, en los casos de réplicas conceptuales (Hunter y Schmidt, 2015), es decir, en los casos en que un estudio presentase más de un indicador de la relación entre los mismos dos constructos para la misma muestra (por ejemplo, dos correlaciones entre capacidad cognitiva y RC, cada una con una medida de capacidad cognitiva diferente), el estudio fue integrado en el meta-análisis siguiendo las siguientes reglas: (a) si el estudio primario presentaba las correlaciones entre las diferentes medidas replicadas se procedió a la acumulación de las medidas en un único resultado mediante el cálculo de la correlación para el compuesto lineal de esas variables. A estos efectos, se siguieron las sugerencias de Hunter y Schmidt (2015) para el cálculo de los compuestos lineales. (b) Si el estudio primario no reportaba las correlaciones entre las medidas replicadas, se integró en el meta-análisis la validez promedio estimada para ese estudio. Aunque esta última estimación es menos precisa que el cálculo del compuesto lineal, ya que subestima el tamaño del efecto real e infla la varianza verdadera (Hunter y Schmidt, 2015), es una buena aproximación del tamaño del efecto observado.
3. En los casos de réplicas conceptuales, si el estudio indicaba que el tamaño de la muestra es distinto para ambas relaciones, es decir, existe mortandad de la muestra, entonces: (a) si la diferencia entre ambos tamaños de la muestra era pequeña, se calculó el compuesto lineal o la validez promedio (según se indicase o no la relación entre las medidas replicadas) y se integró ese tamaño del efecto con el tamaño muestral más pequeño; (b) si la diferencia entre ambos tamaños de la muestra era considerable, se integró únicamente la medida con mayor muestra.
4. Para calcular la fiabilidad de los compuestos lineales se siguieron las siguientes reglas: (a) en los casos en que el estudio primario facilitó un coeficiente de fiabilidad para todas las medidas replicadas se calculó la fiabilidad del compuesto mediante la fórmula de Mosier (1943); (b) si el estudio primario no indicaba el coeficiente de fiabilidad de alguna de las medidas, se tomó la fiabilidad de esa medida reportada en el manual o en otro estudio primario; (c) si no se pudo aplicar la regla anterior se asumió que la fiabilidad de esa medida era la fiabilidad promedio de las restantes medidas agrupadas en el compuesto.

5. Para calcular la fiabilidad de las réplicas conceptuales calculadas mediante la validez promedio (punto 2b), se estimó la fiabilidad a través del promedio de los coeficientes de fiabilidad de las diferentes medidas. En los casos en que el estudio primario no indicase el coeficiente de fiabilidad de alguna de las medidas replicadas, se adoptaron las reglas b y c del punto 4.
6. Una de las medidas de capacidad numérica incluidas en la integración cuantitativa ha sido la *Numeracy Scale* (Cokely et al., 2012; Lipkus et al., 2001; Peters y Levin, 2008; Schwartz et al., 1997; Weller et al., 2013). Estas escalas han sido definidas como medidas que evalúan la capacidad para procesar probabilidades básicas y otros conceptos numéricos, aplicados a resultados del día a día o a contextos de la salud (Cokely et al., 2012; Graffeo, et al., 2015; Peters et al., 2006). No obstante, la gran mayoría de ellas son medidas metodológicamente dudosas, ya que (a) no diferencian adecuadamente entre altos y bajos en capacidad numérica, muestran una importante curtosis negativa, produciendo un efecto “techo” de la medida o existe falta de variabilidad entre las puntuaciones de los sujetos y (b) la fiabilidad de estas escalas tiende a ser baja, con magnitudes de los coeficientes de fiabilidad por consistencia interna en torno a .40 y .60. Por lo tanto, se han llevado a cabo meta-análisis incluyendo y excluyendo estas medidas para controlar sus posibles efectos en los resultados.
7. Durante la fase de recogida de datos, se detectaron tres integraciones sobre las diferencias en función del sexo en la RC (Cueva et al., 2015; Brañas-Garza et al., 2015; Primi et al., 2018). Las integraciones de Cueva et al. (2015) y Brañas-Garza et al. (2015) no fueron incluidas en el meta-análisis por falta de datos para calcular el tamaño del efecto. En su lugar, se procedió a la integración de los estudios primarios a los que se tuvo acceso. La integración de Primi et al. (2018) se incluyó en el meta-análisis, ya que aportaba los datos necesarios para ser integrado y, sin embargo, no se tenía acceso a todos los estudios individuales incluidos en ese meta-análisis. No obstante, con el objetivo de representar la variabilidad encontrada por los autores en la integración ($SD_{\delta} = .095$), se procedió a calcular la distribución de δ partiendo de los datos $\delta = .529$, $SD_{\delta} = .095$, $IC\ 95\%$ ($LI = .34$, $LS = .72$), $K = 13$, $N = 2,536$ facilitados por el estudio. La distribución resultante se introdujo en nuestra integración:
 - a. 1 estudio de $d = .339$ con una $n = 281$ correspondiente a la cola inferior $\bar{X} - 2SD$.
 - b. 2 estudios de $d = .434$ con una $n = 282$ correspondiente a $\bar{X} - 1SD$.

- c. 3 estudios de $d = .529$ con una $n = 282$ correspondiente a $\bar{X} \pm 0.5 SD$.
 - d. 2 estudios de $d = .624$ con una $n = 282$ correspondiente a $\bar{X} + 1SD$.
 - e. 1 estudio de $d = .719$ con una $n = 281$ correspondiente a la cola superior $\bar{X} + 2SD$.
8. En los casos en que un estudio emplease simultáneamente el TRC original y un TRC nuevo, se aplicó la siguiente regla: (a) en el meta-análisis general se integró el tamaño del efecto usando el TRC original ya que, posteriormente, se examinó la variable moderadora tipo de TRC y (b) en el análisis del tipo de TRC como posible variable moderadora se integraron, por un lado, las relaciones con el TRC original y, por otro, las relaciones con nuevos TRC.
 9. Finalmente, se comprobó la dirección de los tamaños del efecto de las diferencias entre hombres y mujeres en RC en función de la codificación asignada a cada grupo (hombres = 1 y mujeres = 0). Del mismo modo, se comprobó la dirección de los tamaños del efecto entre las variables de RC y capacidad cognitiva y desempeño académico. En los casos en que resultó necesario, se unificó el signo de los tamaños del efecto para armonizar los resultados.

3. MÉTODO DE META-ANÁLISIS

Para llevar a cabo los diferentes meta-análisis se emplearon los métodos de meta-análisis de Hunter y Schmidt (2015). Concretamente para examinar la relación entre la RC y la CMG y entre la RC y el desempeño académico se emplearon los métodos de meta-análisis de correlaciones y para integrar las diferencias en función del sexo en la RC se aplicaron los métodos de meta-análisis de efectos experimentales. Los cálculos fueron llevados a cabo mediante el software estadístico desarrollado por Schmidt y Le (2004). Para el cálculo de los intervalos de confianza se empleó el programa VALCOR (Salgado, 1997). Este programa realiza una estimación más precisa de los intervalos de confianza de los tamaños del efecto corregidos al emplear las fórmulas del error de muestreo de las correlaciones corregidas desarrolladas por Bobko y Reick (1980). Frecuentemente, los investigadores aplican a las correlaciones corregidas el nivel de significatividad calculado con el error estándar de los coeficientes de correlación. Sin embargo, los tamaños del efecto corregidos son funciones de los coeficientes de correlación y, por tanto, la magnitud de su error de muestreo es mayor que la de los coeficientes de correlación iniciales o no corregidos (Salgado, 1997). Bobko y Reick (1980) desarrollaron diversas fórmulas para el cómputo del error estándar de las correlaciones corregidas y estas son las aplicadas en el programa VALCOR.

El objetivo del meta-análisis era estimar el tamaño del efecto real o verdadero y para ello es necesario corregir cada uno de los tamaños del efecto observados por las diferentes formas de error (artefactos). Los artefactos considerados en los meta-análisis de correlaciones fueron el error de muestreo, el error de medida en la variable dependiente e independiente y la restricción indirecta en el rango en la variable independiente. Los artefactos considerados en el meta-análisis de efectos experimentales fueron el error de muestreo y el error de medida en la variable dependiente. Como los estudios primarios no ofrecían información acerca de los artefactos para corregir de forma individual cada estudio se crearon distribuciones para cada uno de ellos (véanse los siguientes apartados). Asimismo, para el caso de los meta-análisis de correlaciones se estimó el tamaño del efecto verdadero (ρ), es decir, el tamaño de efecto observado, promedio y ponderado por el tamaño de la muestra corregido por atenuación de la fiabilidad en la medida dependiente e independiente y por restricción indirecta en el rango de la variable independiente; y el tamaño del efecto operativo (r_{op}), es decir, el tamaño del efecto observado, promedio y ponderado por el tamaño de la muestra corregido por atenuación de la fiabilidad en la medida dependiente y por restricción indirecta en el rango de la variable independiente. Finalmente, en los casos de los meta-análisis de efectos experimentales, se estimó el tamaño del efecto verdadero (ρ), que en este caso era el tamaño de efecto observado, promedio y ponderado por el tamaño de la muestra corregido por error de muestreo y por falta de fiabilidad en la medida dependiente.

Distribuciones de la Fiabilidad

Fiabilidad de la Capacidad Cognitiva

Para cada una de las medidas de capacidad cognitiva examinadas, se desarrolló una distribución empírica de los coeficientes de fiabilidad por consistencia interna. De este modo, se creó una distribución para: (1) la CMG, (2) la capacidad verbal, (3) la capacidad numérica, (4) la capacidad mecánico-espacial y (5) la memoria operativa. En el caso de la CMG y la capacidad numérica se crearon dos distribuciones empíricas incluyendo y excluyendo las *Numeracy Scales*, respectivamente. Estas distribuciones se desarrollaron a partir de los coeficientes de fiabilidad por consistencia interna proporcionadas por los estudios singulares, pero previamente se comprobó si dichas muestras se encontraban restringidas en el rango. De ser el caso, los coeficientes de fiabilidad fueron corregidos por restricción en el rango antes de ser incorporados a las distribuciones empíricas. Para la corrección, se empleó la fórmula del Caso I de Thorndike (1949). La Tabla 67 presenta las distribuciones de fiabilidad obtenidas para cada medida de capacidad cognitiva.

Tabla 67.

Distribución de los Coeficientes de Fiabilidad de las Variables de Capacidad Cognitiva

	<i>K</i>	\bar{r}_{xx}	<i>SD</i>	Min. - Máx.
CMG	8	.79	.10	.64 - .88
CMG ^a (incluye NS)	24	.67	.14	.44 - .88
C. Verbal	4	.73	.06	.65 - .80
C. Numérica	1	.72	n.a.	n.a.
C. Numérica ^a (incluye NS)	18	.65	.10	.47 - .82
C. Mecánico-Espacial	1	.79	n.a.	n.a.
Memoria Operativa	1	.79	n.a.	n.a.

Nota. La fiabilidad está expresada mediante el coeficiente de consistencia interna Alfa de Cronbach. *K* = número de casos; \bar{r}_{xx} = coeficiente de fiabilidad promedio; *SD* = desviación típica de la fiabilidad; Min. - Máx. = valores mínimo y máximo de la fiabilidad; CMG = capacidad mental general; NS = *Numeracy Scales*; n.a. = no aplica.

^aIncluye las *Numeracy Scales* como medidas de capacidad cognitiva.

Fiabilidad del Desempeño Académico

Para el caso del desempeño académico, se empleó la distribución de fiabilidad desarrollada por Cuadrado (2018), donde a través de la integración de 7 coeficientes (valor mínimo = .78 y valor máximo = .98) se obtuvo un coeficiente promedio de fiabilidad por consistencia interna de .87 (*SD* = .06).

Fiabilidad del TRC

El coeficiente de fiabilidad del TRC proporcionado por cada estudio ha sido el coeficiente por consistencia interna *KR20*. Como se ha venido indicando a lo largo de esta tesis doctoral, las puntuaciones de TRC se distribuyen de forma sesgada (altos niveles de curtosis y asimetría), lo que provoca dificultades a la hora de aplicar los procesos de análisis bivariados donde se asume la distribución normal de las puntuaciones, como es el caso del cálculo del coeficiente de fiabilidad por consistencia interna. De este modo, el coeficiente por consistencia interna no es el coeficiente más adecuado para estimar la fiabilidad del test.

Por este motivo, para llevar a cabo las correcciones por error de medida del TRC original hemos calculado un coeficiente *CES* a partir de los datos recogidos para esta tesis. Concretamente, hemos seleccionado 3 ítems del TRC-10 (Salgado, 2014a) equivalentes, en cuanto a su contenido, a los 3 ítems originales de Frederick y se han tomado como una forma paralela del TRC original. Como disponemos de las respuestas de los participantes al TRC-3 original y a la forma paralela en dos momentos temporales diferentes hemos calculado el coeficiente *CES* del TRC-3. Para ello: (a) en primer lugar, se ha calculado la correlación tetracórica entre las respuestas del TRC-3

original proporcionadas por los sujetos en un momento temporal 1 y las respuestas a esos mismos sujetos a la versión equivalente en un momento temporal 2; (b) seguidamente, se ha calculado la correlación tetracórica entre las respuestas del TRC-3 original proporcionadas por los sujetos en un momento temporal 2 y las respuestas a esos mismos sujetos a la versión equivalente en un momento temporal 1. El promedio de ambas correlaciones es el coeficiente de fiabilidad *CES* del TRC-3. La magnitud obtenida del coeficiente de fiabilidad *CES* del TRC-3 ha sido .58 ($SD = .048$). No obstante, como la muestra con la que se ha calculado este coeficiente de fiabilidad estaba restringida en el rango, se ha procedido a la corrección del coeficiente *CES* por este artefacto. Para ello se ha empleado la fórmula del Caso I de Thorndike (1949). El resultado final ha sido .63 ($SD = .042$).

El coeficiente de fiabilidad utilizado para corregir por error de medida las relaciones con el TRC-13 (Salgado, 2014a) ha sido el coeficiente *CES* calculado para ese test en el estudio 1 de esta tesis doctoral ($r_{xx} = .66$). En el caso de los estudios en los que se han empleado otras versiones del TRC, se ha hecho una estimación de la magnitud del coeficiente de fiabilidad *CES* de cada una de esas nuevas versiones a través del número de ítems que lo componen. Para ello: (1) en primer lugar, se ha hecho una estimación de la magnitud de la fiabilidad *CES* de cada uno de los ítems del TRC, partiendo de la fiabilidad del TRC-13 y empleando, a tal efecto, la fórmula general de Spearman-Brown:

$$r_{xx} = \frac{n * r_{12}}{1 + (n - 1) * r_{12}}$$

Donde, en nuestro caso, r_{xx} es la fiabilidad *CES* del TRC-13, n es el número de ítems y r_{12} es la estimación del coeficiente *CES* de cada ítem. Esta estimación ha resultado en .13. (2) A continuación, se ha estimado la fiabilidad *CES* de las nuevas versiones a partir de la misma fórmula general de Spearman-Brown y aplicando la estimación de .13 de la fiabilidad *CES* de cada ítem. La Tabla 68 recoge la magnitud del coeficiente de fiabilidad *CES* de las versiones desarrolladas utilizadas en los meta-análisis.

Tabla 68.

Estimación del Coeficiente de Fiabilidad CES de las Nuevas Versiones del TRC según el Número de Ítems

	Estimación de CES
TRC-4 ítems	.37
TRC-5 ítems	.43
TRC-6 ítems	.47
TRC-7 ítems	.51
TRC-11 ítems	.62

Nota. CES = coeficiente de equivalencia y estabilidad

En los meta-análisis donde se examina la variable moderadora “número de ítems” o tipo de TRC, para el TRC-3 se empleó la fiabilidad CES del TRC original y para las nuevas versiones se desarrolló una distribución empírica a partir de los coeficientes de fiabilidad CES calculados en el punto 2. En el caso de los meta-análisis generales se desarrolló una distribución empírica con los coeficientes CES del TRC original y de las versiones desarrolladas incluidas en cada meta-análisis general. Por lo tanto, y en su conjunto, la Tabla 69 muestra la distribución de los coeficientes de fiabilidad de la RC.

Tabla 69.

Distribución de los Coeficientes de Fiabilidad de la RC

	<i>K</i>	\bar{r}_{xx}	<i>SD</i>	Min. - Máx.
<i>Predictor</i>				
RC-DA ^a	15	.59	.07	.43 - .66
RC (TRC-3)	10	.63	n.a.	n.a.
RC (nuevas versiones)	5	.52	.09	.43 - .66
<i>Criterio</i>				
RC-CMG ^b	45	.60	.06	.43 - .63
RC (TRC-3)	37	.63	n.a.	n.a.
RC (nuevas versiones)	12	.50	.06	.43 - .66
RC-Sexo ^c	58	.61	.06	.37 - .66
RC (TRC-3)	50	.63	n.a.	n.a.
RC (nuevas versiones)	13	.46	.09	.37 - .66

Nota. La fiabilidad de la RC está expresada mediante el coeficiente de estabilidad y equivalencia CES. *K* = número de casos; \bar{r}_{xx} = coeficiente de fiabilidad promedio; *SD* = desviación típica de la fiabilidad; Min. - Máx. = valores mínimo y máximo de la fiabilidad; RC = reflexibilidad cognitiva; DA = desempeño académico; TRC-3 = TRC original de Frederick (2005); CMG = capacidad mental general; n.a. = no aplica.

^aRC como variable predictora en el meta-análisis de correlaciones entre RC y desempeño académico.

^bRC como variable criterio en el meta-análisis de correlaciones entre CMG y RC.

^cRC como variable criterio en el meta-análisis de efectos experimentales (diferencias entre hombres y mujeres en RC).

Distribuciones de la Restricción Indirecta en el Rango del Predictor

Se desarrollaron distribuciones de la restricción indirecta en el rango de las variables predictoras: CMG (para el meta-análisis de correlaciones entre CMG y RC) y RC (para el meta-análisis de correlaciones entre RC y desempeño académico). Para el cálculo del coeficiente de restricción en el rango (u) de cada estudio singular (ratio de la desviación típica del grupo restringido (sd) y la desviación típica del grupo no restringido (SD)), se siguió el siguiente procedimiento:

1. Como valor de la desviación típica del grupo restringido (es decir, sd) se tomó la desviación típica de la medida de CMG o RC, según cada caso, reportada por el estudio para esa muestra.
2. Como valor de la desviación típica del grupo no restringido o población (esto es, SD), en el caso de las medidas de CMG se tomó la desviación normativa facilitada por el manual del instrumento o por el artículo donde se publicó la medida. En el caso concreto del instrumento de capacidad numérica *Numeracy Test* de Lipkus et al. (2001) y Peters y Levin (2008) de 11 ítems, se acumularon las desviaciones típicas de cada estudio singular (sd) en una única desviación típica (SD) y esta desviación típica resultante se tomó como la desviación de la población para ese test en concreto. Para llevar a cabo esta acumulación, se partió de las puntuaciones medias y las varianzas de las *Numeracy Scales*, obtenidas por cada muestra integrada, y se empleó la siguiente fórmula para el cálculo de la desviación total de la población:

$$S^2_{\text{Población}} = \bar{X}_{S^2} + S^2_{\bar{X}} \text{ y, por consiguiente, } SD_{\text{Población}} = \sqrt{S^2_{\text{Población}}}$$

Donde, $S^2_{\text{Población}}$ es la varianza acumulada de los diferentes estudios individuales o la varianza de la población, \bar{X}_{S^2} es la media de las varianzas de todos los estudios integrados, $S^2_{\bar{X}}$ es la varianza de las puntuaciones medias y $SD_{\text{Población}}$ es la desviación total de la población. Como resultado, se acumularon 9 SD y se obtuvo una SD de la población de 2.07.

En el caso de la RC, se utilizó la desviación típica acumulada estimada en el meta-análisis que se presentó en el capítulo teórico 2 para el TRC-3. A estos efectos, se siguió el mismo procedimiento de cálculo de la SD de la población empleado en las *Numeracy Scales* de Lipkus et al. (2001) y Peters y Levin (2008). En total, se integraron 113 desviaciones típicas del TRC-3 y se obtuvo una desviación típica poblacional de 1.15. Esta desviación típica acumulada es la mejor aproximación a la SD del TRC-3 de la población. Este procedimiento no se pudo aplicar a

las nuevas versiones del TRC debido a la gran variedad de tests y al limitado número de estudios que emplearon cada uno de ellos. Por ende, no se pudo calcular el coeficiente u de las nuevas versiones del TRC. Únicamente se contó con el coeficiente u del TRC-13.

Tras el cálculo de los coeficientes u de cada estudio individual, se desarrollaron las distribuciones de la restricción en el rango de las variables de capacidad cognitiva y RC, mediante el cálculo de la u promedio y la SD de los coeficientes de restricción. En el caso de la CMG y la capacidad numérica, se desarrollaron dos distribuciones empíricas incluyendo y excluyendo, respectivamente, las *Numeracy Scales*. La Tabla 70 presenta las distribuciones.

Tabla 70.

Distribución de los Coeficientes de Restricción en el Rango (u) de las Variables

	K	\bar{u}	SD	Min. - Máx.
CMG	8	.86	.18	.64 – 1.20
CMG ^a (incluye NS)	16	.87	.16	.64 – 1.20
C. Verbal	-	-	-	-
C. Numérica	-	-	-	-
C. Numérica ^a (incluye NS)	11	.90	.15	.66 – 1.15
C. Mecánico-Espacial	-	-	-	-
Memoria Operativa	3	.84	.08	.76 – .90
RC ^b	9	.91	.09	.74 – 1.01
RC (TRC-3)	9	.91	.09	.74 – 1.01
RC (nuevas versiones)	1	.97	n.a.	n. a.

Nota. K = número de casos; \bar{u} = coeficiente de restricción en el rango promedio; SD = desviación típica de u ; Min. – Máx. = valores mínimo y máximo de u ; CMG = capacidad mental general; NS = *Numeracy Scales*; RC = reflexibilidad cognitiva; TRC-3 = TRC original de Frederick (2005); n.a. = no aplica.

^aIncluyen las *Numeracy Scales* como medidas de capacidad cognitiva.

^bRC como variable predictora del meta-análisis de correlaciones entre RC y desempeño académico.

4. RESULTADOS

Relación de la CMG y la RC

Los resultados de los meta-análisis sobre la relación entre la capacidad cognitiva y la RC se recogen en las Tablas 71 y 72. La Tabla 71 refleja la capacidad de la CMG y de las capacidades cognitivas para explicar la RC sin emplear, a tal efecto, las *Numeracy Scales* como medidas de capacidad numérica y como medidas imperfectas de CMG. La Tabla 72 presenta los resultados de los meta-análisis anteriores repetidos incluyendo las *Numeracy Scales*. Más detalladamente, la Tabla 71 presenta los resultados de 7 meta-análisis. El primero de ellos, se refiere a la relación de la CMG con la RC, con independencia de las medidas de inteligencia y de RC empleadas. El segundo y tercer meta-análisis indican los resultados sobre la relación de la CMG con la RC,

medida esta con el TRC original o con nuevas versiones del TRC, respectivamente. Las siguientes integraciones reflejan los resultados de los meta-análisis de correlaciones entre las capacidades cognitivas específicas y la RC. En estos casos, no se pudo examinar la variable moderadora tipo de TRC por el limitado número de estudios que componen estas integraciones o por el limitado número de estudios desarrollados con las nuevas versiones del test.

La Tabla 72 presenta, en primer lugar, la validez de la CMG con independencia de la media de TRC empleada. Seguidamente muestra los resultados de las integraciones controlando el tipo de TRC utilizado en la evaluación de la RC. Finalmente, se indican los resultados de la validez de la capacidad numérica incluyendo los estudios primarios desarrollados con *Numeracy Scales*.

Las columnas de ambas tablas presentan, de izquierda a derecha: (1) el número de muestras independientes integradas en el análisis (K), (2) el tamaño muestral acumulado (N), (3) el tamaño del efecto observado promedio y ponderado por el tamaño de la muestra (r_{xy}), (4) la desviación típica del tamaño del efecto observado (SD_r), (5) la validez operativa (r_{op}) o, lo que es lo mismo, el tamaño del efecto observado, promedio y ponderado por el tamaño de la muestra corregido por falta de fiabilidad en la medida criterio y por restricción en el rango de la medida predictora, (6) la desviación estándar de la validez operativa (SD_{op}), (7) el tamaño del efecto verdadero (p), es decir, el tamaño del efecto observado corregido por todos los errores artificiales (error de muestreo, error en la medida de la variable criterio y de la variable predictora y restricción en el rango en la variable predictora), (8) la desviación típica del tamaño del efecto verdadero (SD_p), (9) el porcentaje de varianza observada explicada por todos los errores artificiales ($\%VE$); en el caso de la Tabla 72, además, se muestra el porcentaje de la varianza observada que es explicada únicamente por el error de muestreo ($\%VE_{EM}$), (10) el valor mínimo de credibilidad del 90% y (11) los intervalos de confianza del 95% de p .

Los resultados muestran que la CMG determina en gran medida la RC de los individuos. Específicamente, con una muestra total acumulada de 12,490 sujetos, se integraron un total de 45 tamaños del efecto y se obtuvo un tamaño del efecto promedio observado y ponderado por el tamaño de la muestra de .39 ($SD_r = .104$) y una validez operativa de .59 ($SD_{op} = .077$). Una vez corregido el tamaño del efecto promedio observado por todos los errores artificiales (falta de fiabilidad en la medida de las variables predictora y criterio y por restricción en el rango de la variable predictora), se obtuvo una validez verdadera de .65 ($SD_p = .084$). Por lo que la CMG explica el 42.25% de la varianza de la RC. Del total de la varianza observada, el 66.87% es debida a los errores artificiales, por lo que otras variables pueden estar moderando la relación entre los constructos. Asimismo, el valor mínimo de credibilidad del 90% fue .54, por lo que en el 90% de los casos el tamaño del efecto fue igual o superior a .54. Al no contener el cero en el intervalo,

los resultados se generalizan a la población. Además, los límites del intervalo de confianza del 95% fueron .64 y .66, de modo que se rechaza la hipótesis nula y se concluye que la CMG determina de forma considerable la RC de los sujetos. Estos datos apoyan nuestra Hipótesis 1a.

Del mismo modo, los análisis llevados a cabo controlando el tipo de TRC muestran que la CMG es un determinante igualmente válido de la RC con independencia del TRC utilizado. Con una muestra acumulada de 10,349 sujetos, se integraron un total de 38 tamaños del efecto entre CMG y el TRC-3 y se obtuvo un tamaño del efecto verdadero de .63. En el caso de las nuevas versiones del TRC, con una muestra acumulada de 3,754 sujetos, se obtuvo un tamaño del efecto verdadero de .65. En ambos casos, los resultados se generalizaron a la población.

En el caso de los análisis del TRC-3, prácticamente la totalidad de la varianza observada fue explicada por los errores artificiales (96.75%). Sin embargo, en el caso de las otras versiones del TRC, los errores artificiales únicamente explicaron el 31.29% de la varianza observada, lo que sugiere que otras variables moderan esta relación. Posiblemente esta variabilidad sea debida a la diversidad de las nuevas versiones del TRC integradas.

Por otro lado, los resultados también muestran que la capacidad verbal, la capacidad numérica, la capacidad mecánico-espacial y la capacidad de memoria operativa determinan la RC de los individuos ($\rho_{\text{verbal}} = .45$, $\rho_{\text{mecánico-espacial}} = .54$, $\rho_{\text{memoria-operativa}} = .52$ y $\rho_{\text{numérica}} = .74$). En el primer caso, con una muestra total acumulada de 12,490 sujetos, se obtuvo una correlación teórica con la capacidad verbal de .45. Asimismo, los resultados indicaron que la capacidad verbal generaliza su validez a la hora de explicar la RC. No obstante, los estudios integrados en los meta-análisis de capacidad numérica, capacidad mecánico-espacial y de memoria operativa deben ser tomados con precaución, ya que los datos reflejan un error de muestreo de segundo orden. El error de muestreo de segundo orden se produce cuando el número de estudios integrado es limitado. Si un meta-análisis está basado en un amplio número de estudios, la estimación de la media y la desviación estándar contendrá poco error de muestreo. Por el contrario, si el meta-análisis está desarrollado con un número limitado de estudios, las estimaciones meta-analíticas contendrán error de muestreo. A este fenómeno se le denomina error de muestreo de segundo orden (Schmidt y Hunter, 2015). Los resultados de las integraciones de capacidad numérica, capacidad mecánico-espacial y de memoria operativa muestran que el 100% de la varianza observada es explicada por el error de muestreo. Por este motivo, aunque los resultados de estos meta-análisis deben ser tomados con precaución, los datos apoyan nuestras Hipótesis 1b y 1c.

Tabla 71.
Resultados de los Meta-Análisis de la Relación de la Capacidad Cognitiva y la RC, Excluyendo las Numeracy Scales

	<i>K</i>	<i>N</i>	<i>r_{xy}</i>	<i>SD_t</i>	<i>r_{op}</i>	<i>SD_{op}</i>	<i>ρ</i>	<i>SD_ρ</i>	% <i>I/E</i>	<i>VC 90%</i>	<i>IC_ρ 95%</i>
CMG	45	12,490	.39	.104	.59	.077	.65	.084	66.87	.54	.640 - .660
TRC-3	38	10,349	.41	.097	.58	.022	.63	.024	96.75	.60	.618 - .642
TRC-nuevos	11	3,754	.40	.119	.60	.123	.65	.134	31.29	.48	.632 - .668
C. Verbal	13	3,168	.30	.088	.38	.081	.45	.095	48.56	.33	.422 - .478
C. Numérica	3	602	.48	.035	.63	.000	.74	.000	100	.74	.74
C. Mecánico-Espacial	3	624	.39	.029	.48	.000	.54	.000	100	.54	.54
Memoria operativa	11	1,485	.30	.080	.48	.000	.52	.000	100	.52	.52

Nota. *K* = número de casos, *N* = tamaño de la muestra; *r_{xy}* = tamaño del efecto observado promedio y ponderado por el tamaño de la muestra; *SD_t* = desviación estándar de *r_{xy}*; *r_{op}* = validez operativa; *SD_{op}* = desviación estándar de *r_{op}*; *ρ* = validez verdadera; *SD_ρ* = desviación estándar de *ρ*; %*I/E* = porcentaje de varianza explicada por los errores artificiales; *VC90%* = valor de credibilidad del 90%; *IC_ρ 95%* = intervalo de confianza del 95% de *ρ*; CMG = capacidad mental general; TRC-3 = TRC original de Frederick (2005); TRC-nuevos = nuevas versiones del TRC.

Tabla 72.
Resultados de los Meta-Análisis de la Relación de la Capacidad Cognitiva y la RC, Incluyendo las Numeracy Scales

	K	N	r _{xy}	SD _r	r _{op}	SD _{op}	ρ	SD _ρ	%VE _{EM}	%VE _{ART}	VC 90%	IC _p 95%
CMG	58	22,029	.46	.093	.69	.034	.80	.040	18.76	89.86	.75	.795 - .805
TRC-3	49	17,798	.46	.095	.66	.000	.77	.000	19.19	100	.77	.77
TRC-nuevos	14	6,015	.49	.074	.79	.039	.90	.044	24.20	77.95	.84	.895 - .905
C. Numérica	26	13,626	.50	.075	.72	.000	.85	.000	19.19	100	.85	.85

Nota. K = número de casos; N = tamaño de la muestra; r_{xy} = tamaño del efecto observado promedio y ponderado por el tamaño de la muestra; SD_r = desviación estándar de r_{xy}; r_{op} = validez operativa; SD_{op} = desviación estándar de r_{op}; ρ = validez verdadera; SD_ρ = desviación estándar de ρ; %VE_{EM} = porcentaje de varianza observada explicada por el error de muestreo; %VE_{ART} = porcentaje de varianza observada explicada por todos los errores artificiales; VC 90% = valor de credibilidad del 90%; IC_p 95% = intervalo de confianza del 95% de p; CMG = capacidad mental general; TRC-3 = TRC original de Frederick (2005); TRC-nuevos = nuevas versiones del TRC.

Los análisis anteriores fueron repetidos incluyendo las medidas de *Numeracy Scales* como instrumentos de evaluación de la capacidad numérica y como medidas imperfectas de la CMG. Como se puede observar en la Tabla 72, la relación de la CMG y de la capacidad numérica con la RC, incluyendo las *Numeracy Scales*, se incrementa sustancialmente. De acuerdo con los datos, la correlación verdadera con la CMG sería .80 (.77 para predecir el TRC-3 y .90 para predecir otras versiones del test) y con la capacidad numérica sería .85. El hecho de que las *Numeracy Scales* presenten correlaciones observadas moderadas con RC y que los coeficientes de fiabilidad por consistencia interna sean bajos (entre .40 y .60) provocan una estimación al alza del tamaño del efecto verdadero cuando se corrige por los errores artificiales. Si a esto le añadimos que el TRC-3 y muchas de sus nuevas versiones también encajan en la definición de medidas metodológicamente dudosas debido a su limitada capacidad para discriminar entre sujetos reflexivos y no reflexivos, y a su baja fiabilidad, la estimación de la validez verdadera se incrementa sustancialmente. De esta forma, en algunos casos, los resultados muestran que la varianza observada es explicada en su totalidad por los errores artificiales. Como se puede observar, este fenómeno no es tanto debido al error de muestreo (ya que este error explica en torno al 18% y 22% de la varianza total) sino al resto de los errores artificiales (especialmente a la falta de fiabilidad de las medidas). Por lo tanto, los resultados de los meta-análisis que incluyen las *Numeracy Scales* deben ser tomados con precaución. En cualquier caso, estos resultados ponen de manifiesto la necesidad de desarrollar mejores medidas psicométricas de RC y de *numeracy*.

Validez de la RC para predecir el Desempeño Académico

La Tabla 73 refleja los resultados de los meta-análisis de correlaciones entre RC y desempeño académico. La primera integración muestra la validez de la RC para predecir el desempeño académico con independencia de la medida de RC empleada. El segundo y tercer meta-análisis examinan la variable moderadora tipo de TRC.

De izquierda a derecha, las primeras 4 columnas de la tabla presentan el número de muestras independientes integradas en el análisis (K), el tamaño muestral acumulado (N), el tamaño del efecto observado promedio y ponderado por el tamaño de la muestra (r_{xy}) y la desviación típica del tamaño del efecto observado (SD_r). A continuación, las dos siguientes columnas indican la validez operativa (r_{op}), esto es, el tamaño del efecto observado, promedio y ponderado por el tamaño de la muestra atenuado por falta de fiabilidad en la medida criterio y por restricción en el rango de la medida predictora; y la desviación estándar de la validez operativa (SD_{op}). Seguidamente, se presenta el tamaño del efecto verdadero (ρ), es decir, el tamaño del efecto observado corregido por todos los errores artificiales (error de muestreo, error en la

medida de la variable criterio y de la variable predictora y restricción en el rango en la variable predictora); y la desviación típica del tamaño del efecto verdadero (SD_p). Las tres últimas columnas presentan el porcentaje de varianza explicada por los errores artificiales ($\%VE$), el valor mínimo de credibilidad del 90% ($VC\ 90\%$) y los intervalos de confianza del 95% de p ($IC_p\ 95\%$), respectivamente.

Los datos muestran que la RC es un predictor válido del desempeño académico. La validez observada promedio y ponderada por el tamaño de la muestra fue .25 ($SD_r = .11$) y la validez operativa fue .31 ($SD_{op} = .106$). Una vez corregida la validez observada por los errores artificiales de error en la medida de las variables predictora y criterio y por restricción indirecta en el rango de la variable predictora, el valor ascendió a .38 ($SD_p = .128$). Por lo tanto, la RC explica el 14.44% de la varianza del desempeño académico. El valor mínimo de credibilidad del 90% fue .22, lo que indica que en el 90% de los casos la validez de RC para predecir el desempeño académico fue igual o superior a .22. Al no contener el cero, los resultados se pueden generalizar a la población. Asimismo, los datos muestran que el 38.34% de la varianza de la validez observada es debida a los errores artificiales, lo que sugiere que otras variables están moderando la validez de la RC para predecir el desempeño académico. Finalmente, los intervalos de confianza del 95% tampoco contienen el cero, lo que indica que se debe rechazar la hipótesis nula. Los resultados confirman nuestra Hipótesis 2.

Por su parte, los meta-análisis llevados a cabo controlando los efectos del tipo de medida de RC muestran que la validez operativa del TRC-3 para predecir el desempeño académico fue .31 y la de las nuevas versiones desarrolladas del TRC fue .23. Asimismo, la validez verdadera del TRC-3 fue .37 y la de los otros TRC .31. Respectivamente, explican el 13.69% y el 9.61% de la varianza del desempeño académico. El valor mínimo de credibilidad del 90% del TRC-3 fue .20 y el de las otras versiones del test .23. Además, los intervalos de confianza del 95% no contuvieron el 0 en ninguno de los dos casos, de modo que los resultados se pueden generalizar a la población. Finalmente, el porcentaje de varianza observada debida a los errores artificiales fue del 33.77% en el caso del TRC-3 y del 70.69%, en el caso de los nuevos TRC. Estos datos sugieren que otras variables moderan la validez del TRC cuando se emplea la medida original, mientras que, cuando se emplean las nuevas medidas, la mayor parte de los errores artificiales fueron los responsables de explicar la variabilidad de los tamaños del efecto observados. En síntesis, los resultados del análisis del tipo de TRC indican que la medida de RC empleada en la evaluación del constructo afecta a la predicción del desempeño académico, aunque la diferencia en el tamaño del efecto es pequeña.

Tabla 73.
Resultados de los Meta-Análisis de Validez de la RC para Predecir el Desempeño Académico

	<i>K</i>	<i>N</i>	<i>r</i> _{xy}	<i>SD</i> _t	<i>r</i> _{op}	<i>SD</i> _{op}	ρ	<i>SD</i> _ρ	% <i>VE</i>	<i>VC</i> 90%	<i>IC</i> _ρ 95%
RC-DA	11	3,063	.25	.111	.31	.106	.38	.128	38.34	.22	.350 - .410
TRC-3	10	3,007	.25	.112	.31	.110	.37	.132	33.77	.20	.339 - .401
TRC-nuevos	5	1,381	.19	.071	.23	.040	.31	.060	70.69	.23	.262 - .358

Nota. *K* = número de casos; *N* = tamaño de la muestra; *r_{xy}* = tamaño del efecto observado promedio y ponderado por el tamaño de la muestra; *SD_t* = desviación estándar de *r_{xy}*; *r_{op}* = validez operativa; *SD_{op}* = desviación estándar de *r_{op}*; *ρ* = validez verdadera; *SD_ρ* = desviación estándar de *ρ*; *%I/E* = porcentaje de varianza explicada por los errores artificiales; *IC 90%* = valor de credibilidad del 90%; *IC_ρ 95%* = intervalo de confianza del 95% de *ρ*; RC-DA = relación entre la reflexibilidad cognitiva y el desempeño académico; TRC-3 = TRC original de Frederick (2005); TRC-nuevos = nuevas versiones del TRC.

Diferencias entre Hombres y Mujeres en la RC

La Tabla 74 muestra los resultados de los meta-análisis de diferencias en función del sexo en la RC. La primera integración refleja los resultados con independencia del instrumento de RC empleado. El segundo y tercer meta-análisis examinan la variable moderadora tipo de TRC.

De izquierda a derecha, las primeras columnas de la tabla presentan el número de muestras independientes integradas en el análisis (K), el tamaño muestral acumulado (N), el tamaño del efecto observado promedio y ponderado por el tamaño de la muestra (d de cohen o d_{xy}) y la desviación típica del tamaño del efecto observado (SD_d). Seguidamente, se presenta el tamaño del efecto verdadero (δ), es decir, el tamaño del efecto observado y corregido por falta de fiabilidad en la variable dependiente y la desviación típica del tamaño del efecto verdadero (SD_δ). La séptima columna presenta el porcentaje de varianza explicada por los errores artificiales ($\%VE$), la octava refleja el valor mínimo de credibilidad del 90% ($VC\ 90\%$) y la última muestra los intervalos de confianza del 95% de δ ($IC_\delta\ 95\%$).

Los resultados reflejan diferencias entre hombres y mujeres en el desempeño en el TRC. En todos los casos, estas diferencias fueron positivas lo que indica que los hombres alcanzan puntuaciones más altas que las mujeres en RC. Concretamente, en el meta-análisis general se integraron 66 muestras independientes (57 muestras y un meta-análisis que se dividió en 9 muestras). Con un tamaño de la muestra acumulado de 30,529 sujetos, los resultados reflejaron un tamaño del efecto promedio observado de $d = .38$, que una vez corregido por error en la medida de la RC se incrementó a $\delta = .49$. De acuerdo con los criterios de Cohen (1977), el tamaño del efecto se considera medio. El 28.68% de la varianza fue debida a los errores artificiales, lo que sugiere la presencia de otras variables que moderan las diferencias en función del sexo en la RC. Asimismo, el valor mínimo de credibilidad del 90% fue .25, lo que indica que en el 90% de los casos el tamaño del efecto verdadero es igual o superior a .25. Al no incluirse el cero en el intervalo de credibilidad, los resultados se generalizaron a la población. Por su parte, los intervalos de confianza del 95% de δ fueron .47 y .51. Los intervalos de confianza examinan la probabilidad de error en la estimación del tamaño del efecto promedio debido al error de muestreo. De modo que, al no incluirse el cero, se rechaza la hipótesis nula y se puede afirmar que, en efecto, existen diferencias entre hombres y mujeres en RC. Estos resultados apoyan la Hipótesis 3.

Los meta-análisis segundo y tercero examinan las diferencias entre hombres y mujeres en la RC controlando los posibles efectos que el tipo de medida de RC pueda provocar en los resultados. De este modo, se integraron un total de 58 muestras independientes (49 estudios y un meta-análisis que se dividió en 9 muestras) que respondieron al TRC original (Frederick, 2005). Con una muestra acumulada de 27,426 sujetos, el tamaño del efecto observado promedio y ponderado por el tamaño de la muestra fue de .37 que, una vez corregido por el error en la medida de RC, ascendió a .47. El 26.88% de la varianza fue debida a todos los errores artificiales, esto es, el error de muestreo y la falta de fiabilidad en la RC. Nuevamente, el intervalo de credibilidad

excluyó el cero, lo que indica que los resultados son generalizables a la población. Por otro lado, se integraron 13 muestras independientes que respondieron a las nuevas versiones del TRC. La muestra total acumulada fue de 6,062 sujetos y el tamaño del efecto verdadero encontrado fue de .66. En este caso, el 28.35% de la varianza fue debida a los errores artificiales y, nuevamente, el valor mínimo de credibilidad del 90% no incluyó el cero. En base a estos resultados se puede concluir que las diferencias entre hombres y mujeres en RC se incrementan cuando se emplean versiones ampliadas (con más ítems) del TRC.

Tabla 74.

Resultados del Meta-Análisis de Diferencias en Función del Sexo en la RC

	<i>K</i>	<i>N</i>	<i>d_{xy}</i>	<i>SD_d</i>	δ	<i>SD_δ</i>	% <i>VE</i>	<i>VC</i> 90%	<i>IC_δ</i> 95%
RC-Sexo	66 ^a	30,529	.38	.176	.49	.189	28.68	.25	.468 - .512
TRC-3	58 ^b	27,426	.37	.180	.47	.194	26.88	.22	.448 - .492
TRC-nuevos	13	6,062	.45	.177	.66	.211	28.35	.39	.616 - .704

Nota. El signo positivo en los tamaños del efecto indica puntuaciones más altas en RC en los hombres. *K* = número de casos; *N* = tamaño de la muestra; *d_{xy}* = tamaño del efecto observado, promedio y ponderado por el tamaño de la muestra; *SD_d* = desviación estándar de *d_{xy}*; δ = tamaño del efecto corregido; *SD_δ* = desviación estándar de δ ; %*VE* = porcentaje de varianza explicada por los errores artificiales; *VC* 90% = valor de credibilidad del 90%; *IC_δ* 95% = intervalo de confianza del 95% de δ ; RC = reflexibilidad cognitiva; TRC-3 = TRC original de Frederick (2005); TRC-nuevos = nuevas versiones del TRC.

^a9 muestras se corresponden con la distribución del meta-análisis de Primi et al. (2018), por lo que en total serían 57 estudios y 1 meta-análisis (*K* = 58).

^b9 muestras se corresponden con la distribución del meta-análisis de Primi et al. (2018), por lo que en total serían 49 estudios y 1 meta-análisis (*K* = 50).

1. DISCUSIÓN

El principal objetivo de este estudio ha sido ampliar las evidencias de validez de la RC, realizando una revisión meta-analítica sobre la relación de la RC con la capacidad cognitiva y con el desempeño académico y estimando las diferencias entre hombres y mujeres en RC. Este estudio contribuye de diversas formas a la literatura de la RC.

En primer lugar, el estudio presenta el primer meta-análisis de la relación entre RC y capacidad cognitiva. De acuerdo con nuestros resultados, la mejor estimación teórica de la correlación entre la CMG y la RC ha sido .65, de modo que ambas variables comparten un importante porcentaje de varianza (aproximadamente el 42%). Además, se ha demostrado que la relación entre la CMG y la RC se generaliza a la población. Estos resultados apoyan la conclusión de que la RC es un constructo estrechamente ligado con la capacidad cognitiva, aunque no necesariamente se trate de una medida exclusivamente de capacidad cognitiva.

Aunque los estudios primarios se han centrado en examinar la correlación entre la RC y la CMG, nuestro meta-análisis presenta una integración sobre la relación entre la RC y otras

medidas cognitivas. Así, la segunda contribución ha sido demostrar que la capacidad cognitiva, con independencia del tipo de capacidad examinada, explica varianza de la RC y que, además, generaliza su validez. De acuerdo con los datos obtenidos, la correlación verdadera entre la capacidad verbal y la RC ha sido .45, la correlación verdadera con capacidad numérica .74, la correlación con capacidad mecánico-espacial .54 y la correlación con capacidad de la memoria operativa .52. En los tres últimos casos, debido al limitado número de estudios integrados, se ha obtenido un error de muestreo de segundo orden por lo que los datos deben ser tomados con precaución.

La tercera contribución de este estudio ha sido demostrar que la RC es un predictor válido del desempeño académico. La mejor estimación teórica de la validez de la RC para explicar el desempeño académico ha sido .38 y la validez operativa .31. Por lo que, la RC explica el 14.44% de la varianza del desempeño académico. Estos resultados tienen implicaciones prácticas para los procesos de admisión a estudios post-obligatorios (estudios universitarios, ciclos formativos, becas de estudio, etc.) y, consecuentemente, para los procesos de selección de personal. Estudios previos han demostrado que el desempeño académico es un predictor del futuro desempeño ocupacional (Dye y Reck, 1988; Roth et al., 1996; Samson et al., 1984). De este modo, las calificaciones alcanzadas en la etapa educativa tienden a ser empleadas como criterio de selección en las empresas (Roth et al., 1996; Strenze, 2007). Al demostrar que la RC determina el desempeño académico, por extensión se podría asumir que la RC determinará el futuro desempeño ocupacional. Además, los resultados del estudio 3 muestran que la RC determina el desempeño ocupacional. En base a lo expuesto, se sugiere que la RC podría ser usada como una herramienta de toma de decisión en ambos dominios.

Otra de las contribuciones de nuestro estudio ha sido estimar el tamaño del efecto verdadero de las diferencias en la RC en función del sexo. De acuerdo con nuestros resultados, los hombres puntúan .49 desviaciones típicas por encima de las mujeres en RC y estos resultados se generalizan a la población. Aunque, siguiendo los criterios de Cohen, el tamaño del efecto es medio, las diferencias entre ambos grupos aumentan si la RC es evaluada con las nuevas versiones del TRC. Esto sugiere que emplear un test de RC con un número mayor de ítems puede clarificar las diferencias entre hombres y mujeres en el constructo.

A lo largo de esta tesis se ha insistido en las limitaciones psicométricas que afectan al TRC-3. Estas limitaciones ponen en tela de juicio la validez del test para evaluar la RC y, por ello, se ha pretendido examinar las evidencias de validez de la RC teniendo en cuenta el tipo de TRC utilizado. Nuestro estudio ha demostrado que apenas existen diferencias entre los tamaños del efecto teóricos de la CMG y la RC, medida esta con el TRC-3 o con las versiones desarrolladas del mismo. Del mismo modo, existen pocas diferencias entre la validez de la RC, medida con el TRC-3 o con nuevas versiones, en la predicción del desempeño académico. No obstante, estos resultados deben ser tomados con precaución, pues como también se adelantaba en el capítulo 2,

en algunos casos, los test desarrollados con posterioridad siguen estando caracterizados por las mismas limitaciones que el TRC-3. Así pues, algunos de ellos están formados por un número limitado de ítems que mantienen su baja capacidad para discriminar en función del nivel del rasgo latente y producen un efecto “suelo”. Además, poseen bajos niveles de fiabilidad. De este modo, se recomienda el uso de TRC más válidos y fiables como, por ejemplo, el TRC de Salgado (2014a).

Para finalizar se debe hacer mención a las limitaciones presentes en esta investigación. En primer lugar, el número de estudios incluidos en los meta-análisis sobre la relación de la capacidad numérica, capacidad mecánico-espacial y la capacidad de la memoria operativa con la RC es limitado, lo que ha provocado un error de muestreo de segundo orden. En este sentido, se requiere el desarrollo de nuevos estudios singulares que examinen estas relaciones para poder ampliar los análisis meta-analíticos. En segundo lugar, en el caso de los meta-análisis sobre la relación de las capacidades cognitivas específicas y la capacidad de la memoria operativa con RC, no se ha podido analizar la variable tipo de TRC, debido al limitado número de estudios encontrados. Ambas casuísticas han impedido el análisis de variables moderadoras. Finalmente, la tercera limitación ha sido la imposibilidad de examinar la relación de la RC con otros constructos cognitivos, como la velocidad perceptual, por falta de investigación primaria al respecto. Por lo tanto, se requiere más investigación en este sentido.

CONCLUSIONES GENERALES

Esta tesis doctoral se ha llevado a cabo con el objetivo último de profundizar en el conocimiento de la reflexibilidad cognitiva (RC). El desarrollo de este concepto surge de los Modelos Duales de Procesamiento (p. ej., Evans, 1984, 1989, 2006; Stanovich, 1999, 2011). Estos modelos sostienen que el razonamiento humano está guiado por dos tipos de procesamiento cognitivos: uno automático e intuitivo, que opera de forma rápida y con poco o ningún esfuerzo y otro analítico y reflexivo, que opera de forma más lenta y con esfuerzo y concentración. Con el propósito de evaluar este constructo, Frederick (2005; Kahneman y Frederick, 2002) desarrolló el Test de Reflexibilidad Cognitiva (TRC-3). El reducido número de ítems, el poco tiempo necesario para responder al test y su validez aparente han motivado a los investigadores a emplear este test como principal instrumento de medición de RC. Sin embargo, sus limitaciones psicométricas (p. ej., la distribución enormemente asimétrica de sus puntuaciones y la baja fiabilidad) han hecho cuestionarse la robustez de esta medida.

El primer objetivo de esta tesis era analizar las propiedades psicométricas de un nuevo Test de Reflexibilidad Cognitiva con el propósito de: (1) comprobar que esta medida supera las limitaciones del TRC más utilizado (Frederick, 2005) y (2) obtener una medida con una distribución de puntuaciones más robusta (distribución normal) que permita, posteriormente establecer su fiabilidad, validez de constructo y de criterio de modo más preciso. En consecuencia, en el primer estudio se ha examinado la estructura interna del nuevo test (TRC-13; Salgado, 2014a) y se ha comparado con la estructura del TRC-3. Los resultados han mostrado que el nuevo test es una medida más robusta contra las violaciones de la normalidad que ocurren en el TRC-3 y que presenta también mayor fiabilidad.

En primer lugar, se ha podido constatar que las puntuaciones del TRC-13 se distribuyen de forma normal o aproximadamente normal, lo que facilita su uso y la aplicación de análisis estadísticos tradicionales. En segundo lugar, se ha confirmado la unidimensionalidad de la nueva medida de RC, lo que concuerda con estudios previos que han planteado la unidimensionalidad del TRC-3 (Primi et al., 2015). Los resultados del primer estudio han mostrado, también, que el TRC-10 y el TRC-3 comparten un porcentaje sustancial de la varianza. En tercer lugar, se encontró que la fiabilidad de la nueva medida es notablemente más elevada que la del TRC-3 original.

Por tanto, la primera aportación relevante de esta tesis a la literatura de la RC ha sido proporcionar un nuevo test para la medición de la RC, que presenta mejores propiedades psicométricas que el test más ampliamente utilizado (TRC-3). La segunda contribución de esta tesis a la literatura de la RC ha sido estimar la fiabilidad de la medida a través de tres tipos diferentes de coeficientes de fiabilidad, lo que ha permitido establecer el grado de error de medida asociado con las distintas fuentes de variación (Becker, 2000; Cronbach, 1947; Schmidt et al., 2003). Se ha estimado la consistencia interna, la estabilidad temporal y equivalencia y estabilidad del nuevo test. Por ejemplo, el coeficiente de estabilidad temporal y el coeficiente de equivalencia y estabilidad no han sido empleados hasta ahora en el desarrollo y evaluación de los test de medición de la RC. El error transitorio, es decir, las variaciones longitudinales en las respuestas a las medidas que se producen por las variaciones aleatorias en los estados psicológicos de los sujetos a través del tiempo, parece haber sido ignorado en la investigación previa. En este sentido, se trata de una contribución única de esta tesis ya que ni el TRC-3 ni ninguna de las medidas de RC alternativas que se han propuesto en la literatura (p. ej., Primi et al., 2015; Toplak et al., 2014) ha sido examinada desde la perspectiva de las fuentes de error de medida múltiples.

El segundo y tercer objetivos de esta tesis han sido estimar la validez de criterio y constructo de la nueva medida de RC en el ámbito de la Psicología del Trabajo. Desde la publicación del TRC-3, autores de disciplinas diversas (p. ej., psicología cognitiva, economía conductual, psicología diferencial) se han interesado en examinar y validar esta medida. Sin embargo, no se había investigado el concepto de RC y su aplicación en el ámbito de la Psicología del Trabajo. En relación con la validez de criterio, se analizó la capacidad de los tests de RC (es decir, TRC-3; TRC-10 y TRC-13) para predecir el desempeño académico y el desempeño ocupacional. Asimismo, se estimó la validez que la RC añade sobre otras variables individuales en la predicción de ambos criterios. Los resultados mostraron que tales tests de RC predicen el desempeño académico y el desempeño ocupacional y que añaden validez sobre otras medidas de diferencias individuales (p. ej., los test de capacidad mental general y los cuestionarios de personalidad) en la predicción de los resultados académicos y del desempeño ocupacional. Estos hallazgos concuerdan con los encontrados por Salgado et al. (2019). Así, la tercera aportación de esta tesis ha sido mostrar que los test de RC son un predictor válido de criterios académicos y organizacionales y que añaden validez sobre la capacidad conjunta de otras variables de diferencias individuales en la predicción de esos criterios.

En relación con la validez de constructo, se exploró la relación de la RC con capacidad mental general y la personalidad (los cinco factores de personalidad y varias disposiciones de pensamiento). Además, se examinó el porcentaje de varianza que en conjunto esas variables explican de la RC y la contribución que cada una de ellas hace a la explicación de la varianza de la RC. Los resultados indicaron que la RC correlaciona de forma considerable con la capacidad mental general. Asimismo, se ha observado que otras capacidades cognitivas específicas (p. ej.,

la capacidad verbal, la numérica y la mecánico-espacial), así como otros constructos cognitivos (p. ej., memoria operativa) también se relacionan sustancialmente con la RC.

Los resultados han mostrado que los tests de RC también se relacionan con variables de personalidad, aunque la magnitud de la relación es menor que la observada con las capacidades cognitivas. Se ha apreciado una relación directa y de pequeña magnitud con los factores de estabilidad emocional, apertura a la experiencia y conciencia. No obstante, los demás factores de personalidad también se relacionaron con la RC, aunque de forma indirecta a través de las disposiciones de pensamiento. En relación con estas últimas, se encontró que la disposición hacia un pensamiento de atención y concentración en la realización de las tareas, la disposición a asumir riesgos en nuestra vida diaria, la disposición hacia un pensamiento comprometido y de disfrute con el esfuerzo cognitivo y la disposición hacia la independencia de campo, frente a la dependencia de campo, se relacionan directamente con la RC.

Por tanto, la cuarta contribución de esta tesis al conocimiento de la RC ha sido demostrar que la RC es un constructo esencialmente cognitivo (aunque también se relaciona con variables de personalidad) y proponer un modelo de predicción de la RC a partir de las relaciones encontradas en los estudios realizados en esta tesis.

Además de los objetivos mencionados hasta el momento, esta tesis también tuvo como finalidad examinar los efectos de algunos factores relacionados con la aplicación y administración de los tests de RC sobre los resultados y el proceso de respuesta en el test.

Por un lado, se ha investigado si el formato de administración del test (papel y lápiz vs. ordenador) afecta al desempeño en el mismo. Los resultados obtenidos indicaron que se alcanzan mejores resultados (es decir, puntuaciones más elevadas) en el TRC cuando los ítems se presentan en el ordenador. Estos resultados concurren con los encontrados en previos estudios (p. ej., Brañas-Garza et al., 2015). No obstante, los datos sugieren que la modalidad de administración del test (lápiz y papel o computador) únicamente afectaría a los resultados de los tres ítems originales de Frederick (2005), pero no a los nuevos ítems diseñados por Salgado (2014a).

Por otro lado, se ha analizado si el orden de presentación de los ítems afecta a las puntuaciones alcanzadas en el test. Los resultados evidencian que se puede llegar a maximizar los resultados de los participantes si se disponen los tres ítems originales en medio del TRC-13. No obstante, el rendimiento en los nuevos ítems (TRC-10) no cambia significativamente en función de su posición a lo largo del test. Por lo que estos ítems parecen controlar mejor los efectos que produce el orden de presentación de las preguntas en el desempeño en el test. En relación con esto, también se ha examinado si esta variable afecta al tiempo que se invierte en responder al test. Los resultados han mostrado que el orden en que los ítems son administrados no afecta al tiempo de respuesta al TRC-13, aunque sí sugieren que se tiende a emplear significativamente más tiempo en la deliberación de los ítems originales (TRC-3) o de los nuevos ítems (TRC-10) cuando estos son los que encabezan la secuencia de preguntas. Este resultado sugiere que los

primeros ítems pueden ser utilizados como ítems de entrenamiento en la prueba. Asimismo, se ha observado una relación positiva entre el tiempo de respuesta y el resultado en el test. Por lo tanto, estos resultados indicaron que tanto el formato de administración del test como el orden en que se presenten los ítems afectan al desempeño en el TRC-13 y a los resultados en los 3 ítems originales. Este efecto no se había comprobado hasta ahora, lo que constituye una nueva aportación original de esta tesis. También se ha evidenciado que el orden de presentación de los ítems no afecta al tiempo que los participantes invierten en responder al TRC-13. Tampoco esta cuestión había sido examinada previamente, por lo que, también se trata de una contribución exclusiva de esta tesis al conocimiento de la RC, su medición y el funcionamiento de los tests de RC.

Un tercer grupo de objetivos de esta tesis fue examinar las posibles diferencias en función del sexo en la medición de la RC. Los resultados del Estudio 1 y del meta-análisis del Estudio 4 han encontrado diferencias relevantes entre hombres y mujeres en las puntuaciones en el TRC. Se ha observado que, en promedio, los hombres obtienen mejores resultados que las mujeres. En este sentido, los estudios empíricos también han mostrado que:

- (1) Cuando se utilizan nuevas versiones del TRC (con más de tres ítems), las diferencias entre hombres y mujeres se hacen más patentes, lo que sugiere que el número de ítems del test podría ayudar a clarificar/precisar las diferencias entre hombres y mujeres;
- (2) El formato del test (lápiz y papel o computador) y el orden de presentación de los ítems no afectan a las diferencias entre hombres y mujeres en el desempeño en el TRC;
- (3) Los hombres y las mujeres invierten, en general, el mismo tiempo en responder al test, lo que sugiere que las diferencias entre ambos sexos en la RC no determinan el tiempo de respuesta.

En relación con estos resultados, otra aportación de esta tesis ha sido reflejar la existencia de diferencias en la respuesta a los test de RC en función del sexo y estimar el tamaño del efecto de tales diferencias.

Como consecuencia de estos resultados, también se ha investigado si existe validez diferencial en las medidas de RC, de forma que los coeficientes son diferentes en hombres y mujeres. Las principales conclusiones que se han obtenido han sido:

- (1) El nuevo test de RC presenta una distribución normal o prácticamente normal de las puntuaciones en ambos grupos, aunque en el caso de los hombres las puntuaciones presentan un ligero grado de asimetría negativa.
- (2) Los tests de RC son un predictor válido de criterios académicos y organizacionales tanto en hombres como en mujeres.

(3) No existen diferencias en la validez de los tests de RC para predecir el desempeño académico y ocupacional en ambos sexos.

(4) Los determinantes individuales de la RC son los mismos en hombres y mujeres y, en conjunto, explican prácticamente el mismo porcentaje de varianza, aunque la contribución que cada uno de ellos hace es diferente en cada grupo. En el caso de los hombres, las variables individuales que determinan la RC han sido la capacidad mental general, las dimensiones estabilidad emocional, amigabilidad, apertura a la experiencia y conciencia y las disposiciones hacia un pensamiento de autocontrol, de impulsividad, de intuición, de atención en la tarea y de disfrute con la cognición. En el caso de las mujeres, las variables individuales que determinan la RC han sido la capacidad mental general, las dimensiones extraversión, amigabilidad, apertura a la experiencia y conciencia y las disposiciones hacia un pensamiento de impulsividad e intuición o control de los impulsos, de esfuerzo en la tarea, de compromiso y disfrute con la cognición y de no aversión al riesgo.

(5) El orden de presentación de los ítems no afecta al tiempo que las mujeres invierten en responder al test, aunque sí afecta al tiempo que invierten los hombres. Estos tienden a emplear significativamente más tiempo en la deliberación de la respuesta a los ítems que se encuentran posicionados al inicio del test.

(6) La relación entre el tiempo de respuesta y el rendimiento en el TRC es positiva en las mujeres y prácticamente nula en los hombres.

En consecuencia, una aportación única y relevante de la tesis al conocimiento de la RC ha sido mostrar que, aunque existen diferencias entre hombres y mujeres en la RC, tales diferencias no afectan a la validez de los tests de RC para predecir los resultados académicos y el desempeño ocupacional en hombres y mujeres. Por último, esta tesis ha contribuido también al conocimiento de la RC al haber mostrado que los determinantes individuales son prácticamente los mismos en hombres y mujeres, aunque la contribución que cada uno de ellos hace a la explicación de la varianza de la RC es ligeramente diferente en cada sexo.

Implicaciones y Recomendaciones Prácticas

Los resultados de esta tesis doctoral permiten sugerir una serie de implicaciones prácticas para la Psicología del Trabajo y para aquellos ámbitos en los que pueda resultar de utilidad contar con una medición de la RC.

En primer lugar, se recomienda no emplear el TRC-3 como instrumento evaluativo de la RC y sustituirlo por tests metodológicamente más aceptables y mejor validados, como puede ser el TRC-13. Además de presentar mejores niveles de fiabilidad, el hecho de que las puntuaciones en este test se distribuyan de forma normal permite examinar más fácilmente la relación lineal de las puntuaciones con otras variables y emplear coeficientes de correlaciones de la familia de

Pearson, empleados en multitud de análisis multivariantes (p. ej., análisis factorial, regresión múltiple, análisis de ecuaciones estructurales). Todos estos aspectos desembocarán en una mejora de la calidad metodológica de los estudios empíricos que se lleven a cabo con este instrumento.

Por otro lado, los resultados de esta tesis han mostrado que existen varios factores relacionados con la aplicación del TRC que pueden afectar a los resultados en el test. Por ello, el investigador que administre el TRC-13 debe tener esto en consideración y determinar la secuencia y el formato de presentación de los ítems que desee administrar. En todos los casos se debe utilizar la misma secuencia y el mismo formato y no se debe comparar el nivel de RC de los sujetos cuando el test haya sido presentado en un orden o en un formato diferente, especialmente cuando el instrumento se emplee con la finalidad de tomar decisiones que afecten a las personas, como serían los casos de selección de personal. No obstante, si se decide aplicar el TRC-10, es decir, únicamente los nuevos ítems, en este caso no será necesario tener en cuenta las consideraciones realizadas, ya que los datos han indicado que los anteriores factores no afectan a sus resultados.

Esta tesis ha confirmado la existencia de diferencias entre hombres y mujeres en la RC. Por lo tanto, los administradores del test deben aplicar baremos diferentes para ambos grupos para evitar la posibilidad de que se produzca impacto adverso sobre las mujeres en los procesos que impliquen una toma de decisión (Alonso, 2015). En el Anexo IV de este trabajo, se muestran los datos para una interpretación más adecuada de las puntuaciones en el TRC-13 en ambos sexos.

Por último, los resultados también presentan implicaciones específicas para la práctica de selección de personal y para la toma de decisión en procesos académicos. De acuerdo con los resultados, se sugiere que la RC podría ser usada como una herramienta de toma de decisión en ambos contextos. Asimismo, como la RC es un constructo esencialmente cognitivo, en aquellos casos en los que no se tenga acceso a una medida de capacidad mental general, el TRC puede ser utilizado como una medida *proxy* de la misma.

Sugerencias para Futuras Investigaciones

Este trabajo ha permitido profundizar y ampliar el conocimiento sobre el concepto de RC, su medición y su utilidad práctica. No obstante, aún existe cabida para nuevos estudios e investigación. A este respecto, seguidamente se presentan algunas sugerencias para futuras investigaciones que podrían contribuir a incrementar nuestro conocimiento sobre la RC.

En primer lugar, se ha comprobado que el orden de presentación de los ítems mediante ordenador afecta a las respuestas, pero no se sabe si ocurre lo mismo cuando el formato de presentación es, por ejemplo, lápiz y papel. Por ello, un primer estudio sería comprobar si los resultados y conclusiones hallados en esta tesis para la presentación en ordenador se pueden extender a los casos en que el TRC es administrado en lápiz y papel. Asimismo, es conveniente tratar de replicar los resultados del Estudio 2, sobre los efectos del orden de presentación de los

ítems, con otras muestras y en otros países, con el objetivo de comprobar la generalización de nuestros resultados.

En cuanto a la validez de constructo de la RC, una posible sugerencia de investigación sería seguir indagando en la relación de las capacidades cognitivas y la RC. Para ello, se recomienda explorar, no tanto la validez de la capacidad mental general que ha sido más investigada, sino la validez de capacidades cognitivas específicas como la capacidad numérica, la verbal y la capacidad mecánico-espacial, así como la validez de otros constructos cognitivos como la memoria operativa y la velocidad perceptual. El meta-análisis presentado en el Estudio 4 ha puesto de manifiesto la necesidad de más investigación al producirse error de muestreo de segundo orden en algunos de los casos.

Asimismo, serían pertinentes réplicas del Estudio 3, con el objetivo de realizar meta-análisis que permitan estimar de modo más robusto la relación “verdadera” (teórica) entre la RC y las variables de personalidad (los cinco factores de personalidad y las disposiciones de pensamiento). En este sentido, también se recomienda explorar la relación con otras disposiciones de pensamiento diferentes a las examinadas, para obtener un conocimiento más amplio del constructo de RC y poder añadirlas a nuestro modelo explicativo. Además, la muestra de sujetos que han completado el Test de Figuras Enmascaradas (TFE) ha sido pequeña, especialmente en el grupo de hombres, lo que nos ha impedido incluir esta medida en el análisis de regresión. Teniendo en cuenta que los datos han mostrado que presenta una correlación moderada con RC, parece interesante seguir examinando esta relación en posteriores investigaciones.

Por otro lado, se recomienda la realización de más investigaciones sobre la validez de criterio de la RC. Sería conveniente explorar su validez para predecir nuevos criterios académicos y organizacionales y poder establecer la generalización de la validez a través del tipo de criterio. Del mismo modo, se recomienda examinar la validez de la RC para predecir las diferentes dimensiones del desempeño (a saber, desempeño de tarea, desempeño contextual y conductas contraproductivas). Asimismo, sería interesante replicar los resultados del estudio 3 con el propósito de continuar analizando la validez que la RC añade sobre otras variables individuales en la predicción de dichos criterios y comprobar estas evidencias de validez en función del sexo.

Finalmente, otra cuestión que queda pendiente para una mayor investigación está relacionada con las diferencias entre hombres y mujeres en la RC. El meta-análisis de efectos experimentales ha demostrado que los hombres alcanzan sistemáticamente mejores resultados que las mujeres en el TRC. Sin embargo, muy poco se sabe sobre las posibles fuentes que provocan estas diferencias. Nuevos estudios deberían desarrollarse con el propósito de aclarar esta cuestión.

oOo

En definitiva, esta tesis doctoral ha permitido obtener más conocimiento acerca del concepto de RC, de los instrumentos de medición del mismo y de la fiabilidad y validez de dichos instrumentos para que puedan ser utilizados en contextos académicos y organizacionales. No obstante, nuestros hallazgos no agotan las líneas de investigación sobre este tema y todavía queda mucho espacio para futuras investigaciones.



REFERENCIAS

Las referencias con asterisco indican que los estudios han sido incluidos en el meta-análisis.

- Ackerman, P. L. y Heggestad, E. D. (1997). Intelligence, personality, and interests: evidence for overlapping traits. *Psychological Bulletin*, 121, 219-245. doi:10.1037/0033-2909.121.2.219
- *Aczel, B., Bago, B., Szollosi, A., Foldes, A. y Lukacs, B. (2015). Measuring individual differences in decision biases: Methodological considerations. *Frontiers in Psychology*, 6, 1770. doi:10.3389/fpsyg.2015.01770
- Agnew, S. y Harrison, N. (2017). The role of gender, cognitive attributes and personality on willingness to take risks. *Business and Economic Research*, 7, 1-16. doi:10.5296/ber.v7i1.10371
- *Aktaş, B., Yılmaz, O. y Bahçekapılı, H. G. (2017). Moral pluralism on the trolley tracks: Different normative principles are used for different reasons in justifying moral judgments. *Judgment and Decision Making*, 12, 297-307.
- *Albaity, M., Rahman, M. y Shahidul, I. (2014). Cognitive Reflection Test and behavioral biases in Malaysia. *Judgment and Decision Making*, 9, 148-151.
- Albers, W. (2002) Prominence theory as a tool to model boundedly rational decisions. En G. Gigerenzer y R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 297-317). Cambridge: MIT Press.
- Alonso, P. (2015). *Impacto adverso, sesgos y percepciones de los candidatos en las entrevistas de empleo* (Tesis doctoral no publicada). Universidad de Santiago de Compostela, Santiago de Compostela.
- *Alós-Ferrer, C., Garagnani, M. y Hügelshäfer, S. (2016). Cognitive reflection, decision biases, and response times. *Frontiers in Psychology*, 7, 1402. doi:10.3389/fpsyg.2016.01402
- *Alós-Ferrer, C. y Hügelshäfer, S. (2012). Faith in intuition and behavioral biases. *Journal of Economic Behavior & Organization*, 84, 182-192. doi: 10.1016/j.jebo.2012.08.004

- Alter, A. L., Oppenheimer, D. M., Epley, N. y Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, 136, 569-576. doi:10.1037/0096-3445.136.4.569
- Altmann, S. y Falk, A. (2009). *The impact of cooperation defaults on voluntary contributions to public goods* (Working Paper). University of Bonn. doi:10.1.1.319.7302
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education y Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Anderson, C. L., Cullen, A. y Stamoulis, K. (2008). Preference variability along the policy chain in Vietnam. *The Journal of Socio-Economics*, 37, 1729-1745. doi:10.1016/j.soccc.2008.04.011
- Arendasy, M. E. y Sommer, M. (2012). Using automatic item generation to meet the increasing item demands of high-stakes educational and occupational assessment. *Learning and Individual Differences*, 22, 112-117. doi:10.1016/j.lindif.2011.11.005
- Ashton, M. C., Lee, K., Vernon, P. A. y Jang, K. L. (2000). Fluid intelligence, crystallized intelligence, and the openness/intellect factor. *Journal of Research in Personality*, 34, 198-207. doi:10.1006/jrpe.1999.2276
- Baghestanian, S. y Frey, S. (2016). GO figure: Analytic and strategic skills are separable. *Journal of Behavioral and Experimental Economics*, 64, 71-80. doi:10.1016/j.soccc.2015.06.004
- Baghestanian, S., Lugovskyy, V. y Puzzello, D. (2015). Traders' heterogeneity and bubble-crash patterns in experimental asset markets. *Journal of Economic Behavior & Organization*, 117, 82-101. doi:10.1016/j.jebo.2015.06.007
- Baldi, P. L., Iannello, P., Riva, S. y Antonietti, A. (2013). Cognitive reflection and socially biased decisions. *Studia Psychologica*, 55, 265-271. doi:10.21909/sp.2013.04.641
- Bargh, J. A. y Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54, 462-479. doi:10.1037/0003-066X.54.7.462
- Baron, J. (1988). *Thinking and deciding*. New York, USA: Cambridge University Press.
- *Baron, J., Scott, S., Fincher, K. y Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, 4, 265-284. doi:10.1016/j.jarmac.2014.09.003
- Barr, N., Pennycook, G., Stolz, J. A. y Fugelsang, J. A. (2015). Reasoned connections: A dual-process perspective on creative thought. *Thinking & Reasoning*, 21, 61-75. doi:10.1080/13546783.2014.895915

- Barrick, M. R. y Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26. doi:10.1111/j.1744-6570.1991.tb00688.x
- Bates, T. C. y Shieles, A. (2003). Crystallized intelligence as a product of speed and drive for experience: The relationship of inspection time and openness to g and Gc. *Intelligence*, 31, 275-287. doi:10.1016/S0160-2896(02)00176-9
- Bazerman, M. H., Tenbrunsel, A. E. y Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review*, 23, 225-241. doi: 10.5465/amr.1998.533224
- Beatty, E. L. y Thompson, V. A. (2012). Effects of perspective and belief on analytic reasoning in a scientific reasoning task. *Thinking & Reasoning*, 18, 441-460. doi:10.1080/13546783.2012.687892
- Bechara, A. (2004). Disturbances of emotion regulation after focal brain lesions. En R. J. Bradley, *International review of neurobiology*, Vol. 62 (pp. 159-193). California: US. Elsevier academic press.
- Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience*, 8, 1458-1463. doi:10.1038/nn1584
- Becker, G. (2000). How important is transient error in estimating reliability? Going beyond simulation studies. *Psychological Methods*, 5, 370-379. doi:10.1007/BF02289289
- Bennett, G. K. (1969). *Bennett Mechanical Comprehension Test*. San Antonio: Psychological Corporation.
- Bennett, G. K., Seashore, H. G. y Wesman, A. G. (1989). *Differential Aptitude Test*. Marrickville, NSW: Psychological Corporation.
- *Bergman, O., Ellingsen, T., Johannesson, M. y Svensson, C. (2010). Anchoring and cognitive ability. *Economics Letters*, 107, 66-68. doi:10.1016/j.econlet.2009.12.028
- Bertua, C., Anderson, N. y Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational Psychology*, 78, 387-409. doi:10.1348/096317905X26994
- Besedeš, T., Deck, C., Sarangi, S. y Shor, M. (2012). Reducing choice overload without reducing choices. *Review of Economics and Statistics*, 97, 793-802. doi:10.1162/REST_a_00506
- Bigoni, M. y Dragone, D. (2011). *An Experiment on Experimental Instructions* (Working Paper No. 794). Recuperado de SSRN: <https://ssrn.com/abstract=1950545>. doi:10.2139/ssrn.1950545
- Bigoni, M. y Dragone, D. (2012). Effective and efficient experimental instructions. *Economics Letters*, 117, 460-463. doi:10.1016/j.econlet.2012.06.049

- Bobko, P. y Reick, A. (1980). Large sample estimators for standard errors of functions of correlation coefficients. *Applied Psychological Measurement*, 4, 385-398. doi:10.1177/014662168000400309
- Bock, R. D. y Mislevy, R. J. (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied Psychological Measurement*, 6, 431-444. doi:10.1177/014662168200600405
- *Böckenholt, U. (2012). The cognitive-miser response model: Testing for intuitive and deliberate reasoning. *Psychometrika*, 77, 388-399. doi:10.1007/s11336-012-9251-y
- Borghans, L. y Golsteyn, B. H. (2014). Default options and training participation. *Empirical Economics*, 46, 1417-1428. doi:10.1007/s00181-013-0717-2
- *Bosch-Domènech, A., Brañas-Garza, P. y Espín, A. M. (2014). Can exposure to prenatal sex hormones (2D: 4D) predict cognitive reflection? *Psychoneuroendocrinology*, 43, 1-10. doi:10.1016/j.psycueen.2014.01.023
- Bourgeois-Gironde, S. y Van Der Henst, J. B. (2009). How to open the door to System 2: Debiasing the bat-and-ball problem. En S. Watanabe, A. P. Blaisdell, L. Huber y A. Young (Eds.), *Rational animals, irrational humans* (pp. 235-252). Japón: Keio University Press.
- Brainerd, C. J. y Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. En H. W. Reese y R. Kail (Eds.), *Advances in child development and behavior* (pp. 41-100). San Diego, CA: Academic Press.
- *Brañas-Garza, P., García-Munoz, T. y González, R. H. (2012). Cognitive effort in the beauty contest game. *Journal of Economic Behavior & Organization*, 83, 254-260. doi:10.1016/j.jebo.2012.05.018
- Brañas-Garza, P., Kujal, P. y Lenkei, B. (2015). *Cognitive Reflection Test: whom, how, when* (Working Paper, 15-25). Chapman University. Recuperado de: https://digitalcommons.chapman.edu/csi_working_papers/174/
- *Brosnan, M., Hollinworth, M., Antoniadou, K. y Lewton, M. (2014). Is empathizing intuitive and systemizing deliberative? *Personality and Individual Differences*, 66, 39-43. doi:10.1016/j.paid.2014.03.006
- Browne, M. W. (1975). Predictive validity of a linear regression equation. *British Journal of Mathematical and Statistical Psychology*, 28, 79-87. doi:10.1111/j.2044-8317.1975.tb00550.x
- *Browne, M., Pennycook, G., Goodwin, B. y McHenry, M. (2014). Reflective minds and open hearts: Cognitive style and personality predict religiosity and spiritual thinking in a

- community sample. *European Journal of Social Psychology*, 44, 736-742. doi:10.1002/ejsp.2059
- Cacioppo, J. T. y Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131. doi:10.1037/0022-3514.42.1.116
- Calvillo, D. P. y Burgeno, J. N. (2015). Cognitive reflection predicts the acceptance of unfair ultimatum game offers. *Judgment & Decision Making*, 10, 332-341.
- *Campitelli, G. y Gerrans, P. (2014). Does the Cognitive Reflection Test measure cognitive reflection? A mathematical modeling approach. *Memory & Cognition*, 42, 434-447. doi:10.3758/s13421-013-0367-9
- *Campitelli, G. y Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making*, 5, 182-191.
- *Carpenter, J., Graham, M. y Wolf, J. (2013). Cognitive ability and strategic sophistication. *Games and Economic Behavior*, 80, 115-130. doi:10.1016/j.geb.2013.02.012
- Carriedo, N., Elosúa, M. R. y García-Madruga, J. A. (2011). Working memory, text comprehension, and propositional reasoning: A new semantic anaphora WM test. *The Spanish Journal of Psychology*, 14, 37-49. doi:10.5209/rev_SJOP.2011.v14.n1.3
- Cattell, R. B y Cattell, A. K. S. (1986). *Tests de Factor "g", Escalas 2 y 3. Manual* (4ª edición). Madrid: TEA Ediciones, S. A.
- Cesarini, D., Johannesson, M., Magnusson, P. K. y Wallace, B. (2012). The behavioral genetics of behavioral anomalies. *Management Science*, 58, 21-34. doi:10.1287/mnsc.1110.1329
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752-766. doi:10.1037/0022-3514.39.5.752
- Chaiken, S., Liberman, A. y Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. En J. S. Uleman y J. A. Bargh, *Unintended thought* (pp. 212-252). New York: US. The Guilford Press.
- Chaiken, S. y Trope, Y. (1999). *Dual-process theories in social psychology*. New York, EE. UU: Guilford Press.
- Chandler, J., Mueller, P. y Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, 46, 112-130. doi:10.3758/s13428-013-0365-7
- Cheyne, J. A. y Pennycook, G. (2013). Sleep paralysis postepisode distress: Modeling potential effects of episode characteristics, general psychological distress, beliefs, and cognitive style. *Clinical Psychological Science*, 1, 135-148. doi:10.1177/2167702612466656

- Coates, E. y Blaszczynski, A. (2014). Predictors of return rate discrimination in slot machine play. *Journal of Gambling Studies*, 30, 669-683. doi:10.1007/s10899-013-9375-8
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, J. y Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. T., Scotland, E. y Wolfe, D. M. (1955). An experimental investigation of need for cognition. *Journal of Abnormal and Social Psychology*, 51, 291-294. doi:10.1037/h0042761
- *Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S. y Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, 7, 25-47.
- *Cokely, E. T. y Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, 4, 20-33.
- Conger, A. J. (1974). A revised definition of suppressor variables: A guide to their identification and interpretation. *Educational and Psychological Measurement*, 34, 35-46. doi:10.1177/001316447403400105
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K. y Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology*, 89, 469-487. doi:10.1037/0022-3514.89.4.469
- *Corgnet, B., Espín, A. M. y Hernán-González, R. (2015). Creativity and cognitive skills among millennials: Thinking too much and creating too little. *Frontiers in Psychology*, 7, 1626. doi:10.3389/fpsyg.2016.01626
- *Corgnet, B., Espín, A. M. y Hernán-González, R. (2016). The cognitive basis of social behavior: Cognitive reflection overrides antisocial but not always prosocial motives. *Frontiers in Behavioral Neuroscience*, 9, 287. doi:10.3389/fnbeh.2015.00287
- Corgnet, B., Espín, A. M., Hernán-González, R., Kujal, P. y Rassinetti, S. (2015). To trust, or not to trust: Cognitive reflection in trust games. *Journal of Behavioral and Experimental Economics*, 64, 20-27. doi:10.1016/j.socec.2015.09.008
- *Corgnet, B., Hernán-Gonzalez, R. H. y Mateo, R. (2015). Cognitive reflection and the diligent worker: An experimental study of millennials. *PloS one*, 10, e0141243. doi:10.1371/journal.pone.0141243

- Costa, A., Foucart, A., Arnon, I., Aparici, M. y Apesteguia, J. (2014). “Piensa” twice: On the foreign language effect in decision making. *Cognition*, 130, 236-254. doi:10.1016/j.cognition.2013.11.010
- Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J. y Keysar, B. (2014). Your morals depend on language. *PloS one*, 9, e94842. doi:10.1371/journal.pone.0094842
- Costa, P. T. y McCrae, R. R. (1988). From catalog to classification: Murray’s needs and the five-factor model. *Journal of Personality and Social Psychology*, 55, 258–265. doi:10.1037/0022-3514.55.2.258
- Cronbach, L. J. (1947). Test “reliability”: Its meaning and determination. *Psychometrika*, 12, 1-16. doi:10.1037/1082-989X.5.3.370
- Cuadrado, D. (2018). *Deshonestedad académica, desempeño y diferencias individuales* (Tesis doctoral no publicada). Universidad de Santiago de Compostela, Santiago de Compostela.
- Cueva, C., Iturbe-Ormaetxe, I., Mata-Pérez, E., Ponti, G., Sartarelli, M., Yu, H. y Zhukova, V. (2015). Cognitive (ir) reflection: New experimental evidence. *Journal of Behavioral and Experimental Economics*, 64, 81-93. doi:10.1016/j.socec.2015.09.002
- De Neys, W., d’Ydewalle, G., Schacken, W. y Vos, G. (2002). A Dutch, computerized, and group administrable adaptation of the operation span test. *Psychologica Belgica*, 42, 177-190.
- De Neys, W., Rossi, S. y Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, 20, 269-273. doi:10.3758/s13423-013-0384-5
- *Del Missier, F., Mäntylä, T. y De Bruin, W. B. (2011). Decision-making competence, executive functioning, and general cognitive abilities. *Journal of Behavioral Decision Making*, 25, 331-351. doi:10.1002/bdm.731
- Del Missier, F., Visentini, M. y Mäntylä, T. (2015). Option generation in decision making: Ideation beyond memory retrieval. *Frontiers in Psychology*, 5, 1584. doi:10.3389/fpsyg.2014.01584
- Deppe, K. D., Gonzalez, F. J., Neiman, J. L., Jacobs, C., Pahlke, J., Smith, K. B. y Hibbing, J. R. (2015). Reflective liberals and intuitive conservatives: A look at the Cognitive Reflection Test and ideology. *Judgment & Decision Making*, 10, 314-331.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440. doi:10.1146/annurev.ps.41.020190.002221
- Dodrill, C. B. (1983). Long-term reliability of the Wonderlic Personnel Test. *Journal of consulting and clinical psychology*, 51, 316-317. doi:10.1037/0022-006X.51.2.316

- *Drummond, C. y Fischhoff, B. (2017). Development and validation of the scientific reasoning scale. *Journal of Behavioral Decision Making*, 30, 26-38. doi:10.1002/bdm.1906
- Duttie, K. (2015). Cognitive skills and confidence: Interrelations with overestimation, overplacement and overprecision. *Bulletin of Economic Research*, 68, 42-55. doi:10.1111/boer.12069
- *Duttie, K. e Inukai, K. (2015). Complexity aversion: Influences of cognitive abilities, culture and system of thought. *Economic Bulletin*, 35, 846-855.
- Dye, D. A. y Reck, M. (1988). *A literature review and meta-analysis of education as a predictor of job performance* (OPRD-88-9). Washington, DC: Office of Personnel Management.
- Ekstrom, R. B., Dermen, D. y Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests* (Vol. 102). Princeton, NJ: Educational testing service.
- Elqayam, S. (2009). Models of dependence and independence: A two-dimensional architecture of dual processing. *Thinking & Reasoning*, 15, 377-387. doi:10.1080/13546780903268895
- Else-Quest, N. M., Hyde, J. S. y Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136, 103-127. doi:10.1037/a0018053
- Epstein, S. (2003). Cognitive-experiential self-theory of personality. En T. Millon y M. J. Lerner (Eds.), *Comprehensive handbook of psychology*, Vol. 5, (pp. 159-184). New Jersey, EE.UU: John Wiley y Sons, Inc.
- Epstein, S., Pacini, R., Denes-Raj, V. y Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, 71, 390-405.
- Eriksson, K. y Jansson, F. (2016). Procedural priming of a numerical cognitive illusion. *Judgment and Decision Making*, 11, 205-212.
- Evans, J. S. B. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology*, 75, 451-468. doi: 10.1111/j.2044-8295.1984.tb01915.x
- Evans, J. S. B. (1989). *Bias in human reasoning. Causes and consequences*. London: U.K: Erlbaum.
- Evans, J. S. B. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, 13, 378-395. doi:10.3758/BF03193858
- Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278. doi:10.1146/annurev.psych.59.103006.093629

- Evans, J. S. B. (2009). How many dual-process theories do we need? One, two, or many? En J. S. B. Evans y K. Frankish (Eds.), *In two minds: Dual processes and beyond*, (pp. 55-33). New York, UU.EE: Oxford University Press.
- Evans, J. S. B. y Over, D. E. (1996). *Rationality and reasoning*. Erlbaum: UK. Psychology Press.
- Evans, J. S. B., Over, D. E. y Handley, S. J. (2003). A theory of hypothetical thinking. En D. Hardman y L. Macchi (Eds.), *Thinking: psychological perspectives on reasoning, judgment and decision making*, (pp. 3-22). England: John Wiley & Sons, LTD.
- Evans, J. S. B. y Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8, 223-241. doi:10.1177/1745691612460685
- Evans, J. S. B. y Wason, P. C. (1976). Rationalization in a reasoning task. *British Journal of Psychology*, 67, 479-486. doi:10.1111/j.2044-8295.1976.tb01536.x
- Fairley, K. y Weitzel, U. (2017). Ambiguity and risk measures in the lab and students' real-life borrowing behavior. *Journal of Behavioral and Experimental Economics*, 67, 85-98. doi:10.1016/j.socec.2016.12.001
- *Fehr, D. y Huck, S. (2015). Who knows it is a game? On strategic awareness and cognitive ability. *Experimental Economics*, 19, 713-726. doi:10.1007/s10683-015-9461-0
- Feltz, A. y Cokely, E. T. (2008). The fragmented folk: More evidence of stable individual differences in moral judgments and folk intuitions. En B.C. Love, K. McRae y V.M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 1771-1776). Austin, TX: Cognitive Science Society
- Fernbach, P. M., Sloman, S. A., Louis, R. S. y Shube, J. N. (2013). Explanation fiends and foes: How mechanistic detail determines understanding and preference. *Journal of Consumer Research*, 39, 1115-1131. doi:10.1086/667782
- Ferrando, P. J. y Lorenzo-Seva, U. (2014). Exploratory item factor analysis: Additional considerations. *Anales de Psicología*, 30, 1170-1175. doi:10.6018/analesps.30.3.199991
- Ferrando, P. J. y Lorenzo-Seva, U. (2016). A note on improving EAP trait estimation in oblique factor-analytic and item response theory models. *Psicológica*, 37, 235-247.
- Ferrando, P. J. y Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema*, 29, 236-240. doi: 10.7334/psicothema2016.304
- Ferrara, M., Bottasso, A., Tempesta, D., Carrieri, M., De Gennaro, L. y Ponti, G. (2015). Gender differences in sleep deprivation effects on risk and inequality aversion: Evidence from an economic experiment. *PloS one*, 10, e0120029. doi:10.1371/journal.pone.0120029
- *Finucane, M. L. y Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and Aging*, 25, 271-288. doi:10.1037/a0019106

- Fiske, S. T. y Taylor, S. E. (1988). *Social cognition*. New York, USA: McGraw-Hill Book Company.
- Fosgaard, T. R., Hansen, L. G. y Piovesan, M. (2013). Separating Will from Grace: An experiment on conformity and awareness in cheating. *Journal of Economic Behavior & Organization*, 93, 279-284. doi:10.1016/j.jebo.2013.03.027
- Frankish, K. (2010). Dual-process and dual-system theories of reasoning. *Philosophy Compass*, 5, 914-926. doi:10.1111/j.1747-9991.2010.00330.x
- *Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19, 25-42. doi:10.1257/089533005775196732
- Friedl, A., Neyse, L. y Schmidt, U. (2018). Payment scheme changes and effort adjustment: The role of 2D: 4D digit ratio. *Journal of Behavioral and Experimental Economics*, 72, 86-94. doi:10.1016/j.soccc.2017.11.007
- Fritz, M. S., Kenny, D. A. y MacKinnon, D. P. (2016). The combined effects of measurement error and omitting confounders in the single-mediator model. *Multivariate Behavioral Research*, 51, 681-697. doi:10.1080/00273171.2016.1224154
- Gangemi, A., Bourgeois-Gironde, S. y Mancini, F. (2015). Feelings of error in reasoning—in search of a phenomenon. *Thinking & Reasoning*, 21, 383-396. doi:10.1080/13546783.2014.980755
- Gawronski, B. y Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692-731. doi:10.1037/0033-2909.132.5.692
- Gervais, W. M. (2015). Override the controversy: Analytic thinking predicts endorsement of evolution. *Cognition*, 142, 312-321. doi:10.1016/j.cognition.2015.05.011
- Ghazal, S., Cokely, E. T. y Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making*, 9, 15-34.
- *Gino, F. y Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, 102, 445-459. doi:10.1037/a0026406
- Glöckner, A. y Witteman, C. (2010). Beyond dual-process models: A categorisation of processes underlying intuitive judgement and decision making. *Thinking & Reasoning*, 16, 1-25. doi:10.1080/13546780903395748
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.
- *Goldwater, M. B., Don, H. J., Krusche, M. J. y Livesey, E. J. (2018). Relational discovery in category learning. *Journal of Experimental Psychology: General*, 147, 1-33. doi:10.1037/xge0000387

- Gómez-Chacón, I., García Madruga, J. A., Rodríguez, R., Vila, J. O. y Elosúa, R. (2011). Mathematical beliefs and cognitive reflection: Do they predict academic achievement? En B. Roesken y M. Casper (Eds.), *Current state of research on mathematical beliefs XVII, proceedings of the MAVI-17 conference* (pp. 64–73). Bochum, Germany: Professional School of Education, Ruhr- Universität Bochum
- *Gómez-Chacón, I. M., García-Madruga, J. A., Vila, J. Ó., Elosúa, M. R. y Rodríguez, R. (2014). The dual processes hypothesis in mathematics performance: Beliefs, cognitive reflection, working memory and reasoning. *Learning and Individual Differences*, 29, 67-73. doi:10.1016/j.lindif.2013.10.001
- Goodman, J. K., Cryder, C. E. y Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, 26, 213-224. doi:10.1002/bdm.1753
- *Graffeo, M., Polonio, L. y Bonini, N. (2015). Individual differences in competent consumer choice: The role of cognitive reflection and numeracy skills. *Frontiers in Psychology*, 6, 844. doi:10.3389/fpsyg.2015.00844
- Gronwall, D. M. A. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367–373. doi:10.2466/pms.1977.44.2.367
- *Grossman, Z., van der Weele, J. y Andrijevik, A. (2014). *A test of dual-process reasoning in charitable giving* (Working Paper). University of California Santa Bárbara. Recuperado de: <https://escholarship.org/uc/item/4tm617f7>
- Guilford, J. P. y Fruchter, B. (1984). *Estadística aplicada a la psicología y la educación*. México: McGraw-Hill.
- Guillen, P. y Hakimov, R. (2014). Monkey see, monkey do: Truth-telling in matching algorithms and the manipulation of others (Discussion Paper No. SP II 2014-202). Recuperado de: <http://hdl.handle.net/10419/95868>
- *Guthrie, C., Rachlinski, J. J. y Wistrich, A. J. (2007). Blinking on the bench: How judges decide cases. *Cornell Law Review*, 93, 1-44.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814-834. doi:10.1037/0033-295X.108.4.814
- Haidt, J. (2008). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. En J. E. Adler y L. J. Rips (Eds.), *Reasoning. studies of human inference and its foundations* (pp. 1024–1053). Cambridge, U.K: Cambridge University Press.
- Hammond, K. R. (1996). *Human judgment and social policy*. New York: US. Oxford University Press.

- *Hanaki, N., Jacquemet, N., Luchini, S. y Zylbersztejn, A. (2016). Fluid intelligence and cognitive reflection in a strategic environment: Evidence from dominance-solvable games. *Frontiers in Psychology*, 7, 1188. doi:10.3389/fpsyg.2016.01188
- Haran, U., Ritov, I. y Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Judgment and Decision Making*, 8, 188-201.
- Hardisty, D. J. y Weber, E. U. (2009). Discounting future green: Money versus the environment. *Journal of Experimental Psychology: General*, 138, 329-340. doi:10.1037/a0016433
- Hogan, R. (1991). Personality and personality measurement. En M. D. Dunnette y L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, págs. 327–396). Palo Alto, CA: Consulting Psychologists Press.
- Hopkins, E. J., Weisberg, D. S. y Taylor, J. C. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition*, 155, 67-76. doi:10.1016/j.cognition.2016.06.011
- Hoppe, E. I. y Kusterer, D. J. (2011). Behavioral biases and cognitive reflection. *Economics Letters*, 110, 97-100. doi:10.1016/j.econlet.2010.11.015
- Horst P. (1941). The role of predictor variables which are independent of the criterion. *Social Science Research Council*. 48, 431-436.
- Howarth, S., Handley, S. J. y Walsh, C. (2016). The logic-bias effect: The role of effortful processing in the resolution of belief–logic conflict. *Memory & Cognition*, 44, 330-349. doi: 10.3758/s13421-015-0555-x
- Huang, M. H. y Hauser, R. M. (1998). Trends in black–white test score differentials: II the WORDSUM Vocabulary Test. En U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 303–332). Washington, DC: American Psychological Association.
- Hughes, T. M y Douzenis, C. (1986). Predictor and performance variables in a performance-based education course. *The Journal of Psychology*, 120, 143-147. doi:10.1080/00223980.1986.9712623
- Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340–362. doi:10.1016/0001-8791(86)90013-8
- Hunter, J. E. y Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98. doi:10.1037/0033-2909.96.1.72
- Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A. y Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly*, 14, 299-324. doi:10.1111/j.1471-6402.1990.tb00022.x

- Ibanez, M., Riener, G. y Rai, A. (2013). *Sorting through affirmative action: Two field experiments in Colombia* (Working Paper N° 150). Courant Research Centre: Poverty, Equity and Growth.
- Inquisit (Version 5.0.6.0) [Software de ordenador]. (2016) Seattle, WA: Millisecond Software.
- Insler, M., Compton, J. y Schmitt, P. (2013). Does everyone accept a free lunch? Decision-making under (almost) zero-cost borrowing. En S. M. Collins, R. M. Isaac, D. A. Norton (Eds.), *Experiments in financial economics* (pp. 145-170). Bingley, UK: Emerald Group Publishing Limited.
- *Insler, M., Compton, J. y Schmitt, P. (2015). The investment decisions of young adults under relaxed borrowing constraints. *Journal of Behavioral and Experimental Economics*, 64, 106-121. doi:10.1016/j.soccc.2015.07.004
- Iyer, R., Koleva, S., Graham, J., Ditto, P. y Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PloS one*, 7, e42366. doi:10.1371/journal.pone.0042366
- *Jackson, S. A., Kleitman, S., Howie, P. y Stankov, L. (2016). Cognitive abilities, monitoring confidence, and control thresholds explain individual differences in heuristics and biases. *Frontiers in Psychology*, 7, 1559. doi:10.3389/fpsyg.2016.01559
- Johnson, E. D., Tubau, E. y De Neys, W. (2014). The unbearable burden of executive load on cognitive reflection: A validation of dual process theory. En *Proceedings of the annual meeting of the Cognitive Science Society*: Vol. 36, No. 36 (pp. 2441-2446). Recuperado de: <https://cloudfront.escholarship.org/dist/prd/content/qt01x9w1hw/qt01x9w1hw.pdf>
- Jöreskog, K. y Sörbom, D. (1998). LISREL 8.20 and PRELIS 2.20 [programa informático]. Chicago: Scientific Software Inc.
- Juanchich, M., Dewberry, C., Sirota, M. y Narendran, S. (2016). Cognitive reflection predicts real-life decision outcomes, but not over and above personality and decision-making styles. *Journal of Behavioral Decision Making*, 29, 52-59. doi:10.1002/bdm.1875
- Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, 8, 407-424. doi:10.2139/ssrn.2182588
- Kahneman, D. (2011). *Pensar rápido, pensar despacio*. Barcelona, España: Debolsillo.
- Kahneman, D. y Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. En T. Gilovich, D. Griffin, y D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York, US: Cambridge University Press.

- Kahneman, D. y Frederick, S. (2005). A model of heuristic judgment. En K. J. Holyoak y R. G. Morrison (Eds), *The Cambridge handbook of thinking and reasoning* (pp. 267-293). New York, EE. UU: Cambridge University Press.
- Kamei, K. (2016). Democracy and resilient pro-social behavioral change: an experimental study. *Social Choice and Welfare*, 47, 359-378. doi:10.1007/s00355-016-0967-y
- Kanoy, K. W., Wester, J. L. y Latta, M. (1989). Predicting college success of freshmen using traditional, cognitive, and psychological measures. *Journal of Research and Development in Education*, 22, 65-70.
- Kinnunen, S. P. y Windmann, S. (2013). Dual-processing altruism. *Frontiers in Psychology*, 4, 193. doi:10.3389/fpsyg.2013.00193
- *Kiss, H. J., Rodriguez-Lara, I. y Rosa-García, A. (2016). Think twice before running! Bank runs and cognitive abilities. *Journal of Behavioral and Experimental Economics*, 64, 12-19. doi:10.1016/j.socec.2015.01.006
- Knežević, G. y Opačić, G. (2011). *Vocabulary Test*. Unpublished material.
- Kocher, M., Lucks, K. E. y Schindler, D. (2016). *Unleashing animal spirits-self-control and overpricing in experimental asset markets* (Working Paper No. 2016-3). Recuperado de: <https://ssrn.com/abstract=3140254>
- Koehler, D. J. y James, G. (2010). Probability matching and strategy availability. *Memory & Cognition*, 38, 667-676. doi:10.3758/MC.38.6.667
- Königsheim, C., Lukas, M. y Nöth, M. (2018). Individual preferences and the exponential growth bias. *Journal of Economic Behavior & Organization*, 145, 352-369. doi:10.1016/j.jebo.2017.07.032
- Koriat, A. (2017). Can people identify “deceptive” or “misleading” items that tend to produce mostly wrong answers?. *Journal of Behavioral Decision Making*, 30, 1066-1077. doi:10.1002/bdm.2024
- Koscielniak, M., Rydzewska, K. y Sedek, G. (2016). Effects of age and initial risk perception on Balloon Analog Risk Task: The mediating role of processing speed and need for cognitive closure. *Frontiers in Psychology*, 7, 659. doi:10.3389/fpsyg.2016.00659
- Kramer, M. M. (2016). Financial literacy, confidence and financial advice seeking. *Journal of Economic Behavior & Organization*, 131, 198-217. doi:10.1016/j.jebo.2016.08.016
- Krawczyk, M. y Sylwestrzak, M. (2018). Exploring the role of deliberation time in non-selfish behavior: The double response method. *Journal of Behavioral and Experimental Economics*, 72, 121-134. doi:10.1016/j.socec.2017.12.004
- Kruglanski, A. W. y Gigerenzer, G. (2011). Intuitive and deliberative judgments are based on common principles. *Psychological Review*, 118, 97-109. doi:10.1037/a0020762

- Kuhn, M., Kuhn, P. y Villeval, M. C. (2014). *Self control and intertemporal choice: Evidence from glucose and depletion interventions* (Working Paper Series No. 4609). Recuperado de: <https://ssrn.com/abstract=2398279>
- Kuhn, M., Kuhn, P. y Villeval, M. C. (2017). Decision-environment effects on intertemporal financial choices: How relevant are resource-depletion models? *Journal of Economic Behavior & Organization*, 137, 72-89. doi:10.1016/j.jebo.2017.02.014
- Kuncel, N. R. (2003). *The prediction and structure of academic performance*. Tesis doctoral no publicada. University of Minnesota, Minneapolis, MN.
- Kuncel, N. R., Credé, M. y Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, 75, 63-82. doi:10.3102/00346543075001063
- Kuncel, N. R., Hezlett, S. H. y Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127, 162-181. doi:10.1037//Q033-2909.127.1.162
- Langnickel, F. y Zeisberger, S. (2016). Do we measure overconfidence? A closer look at the interval production task. *Journal of Economic Behavior & Organization*, 128, 121-133. doi:10.1016/j.jebo.2016.04.019
- Lautenschlager, G. J. (1990). Sources of imprecision in formula cross-validated multiple correlations. *Journal of Applied Psychology*, 75, 460-462. doi:10.1037/0021-9010.75.4.460
- Lawson, J., Baron-Cohen, S. y Wheelwright, S. (2004). Empathising and systemising in adults with and without asperger syndrome. *Journal of Autism and Developmental Disorders*, 34, 301-310. doi:10.1023/B:JADD.0000029552.42724.1b.
- Lesage, E., Navarrete, G. y De Neys, W. (2013). Evolutionary modules and bayesian facilitation: The role of general cognitive resources. *Thinking & Reasoning*, 19, 27-53. doi:10.1080/13546783.2012.713177
- Leykin, Y. y DeRubeis, R. J. (2010). Decision-making styles and depressive symptomatology: Development of the Decision Styles Questionnaire. *Judgment and Decision Making*, 5, 506-515.
- *Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M. y Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, 25, 361-381. doi:10.1002/bdm.752

- Lieberman, M. D. (2003). Reflexive and reflective judgment processes. A social cognitive neuroscience approach. En J. P. Forgas, K. D. Williams y W. von Hippel (Eds.), *Social judgment. Implicit and explicit processes*, (pp. 44-67). Melbourne, Australia: Cambridge University Press.
- Lindberg, S. M., Hyde, J. S., Petersen, J. L. y Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, 136, 1123-1135. doi:10.1037/a0021276
- *Lindeman, M. y Svedholm-Häkkinen, A. M. (2016). Does poor understanding of physical world predict religious and paranormal beliefs? *Applied Cognitive Psychology*, 30, 736-742. doi:10.1002/acp.3248
- Lindenberger, U., Mayr, U. y Kliegl, R. (1993). Speed and intelligence in old age. *Psychology and Aging*, 8, 207-220. doi:10.1037/0882-7974.8.2.207
- Lipkus, I. M., Samsa, G. y Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21, 37-44. doi:10.1177/0272989X0102100105
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30, 1151-1169. doi:10.6018/analesps.30.3.199361
- Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65, 272-292. doi:10.1006/obhd.1996.0028
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492-527. doi:10.1037/0033-295X.95.4.492
- *Lohse, J. (2016). Smart or selfish—when smart guys finish nice. *Journal of Behavioral and Experimental Economics*, 64, 28-40. doi:10.1016/j.socec.2016.04.002
- Lohse, T., Simon, S. A. y Konrad, K. A. (2018). Deception under time pressure: Conscious decision or a problem of awareness? *Journal of Economic Behavior & Organization*, 146, 31-42. doi:10.1016/j.jebo.2017.11.026
- Lubian, D. y Untertrifaller, A. (2013). *Cognitive ability, stereotypes and gender segregation in the workplace* (Working Paper No. 25/2013). University of Verona. Recuperado de: <http://dse.univr.it/workingpapers/wp2013n25.pdf>
- Macchi, L. y Bagassi, M. (2012). Intuitive and analytical processes in insight problem solving: A psycho-rhetorical approach to the study of reasoning. *Mind & Society*, 11, 53-67. doi:10.1007/s11299-012-0103-3
- MacKinnon, D. (2008). *Introduction to statistical mediation analysis*. New York: Lawrence Erlbaum Associates.

- MacKinnon, D. P., Lockwood, C. M. y Williams, J. (2007). Confidence limits for indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39, 99-128. doi:10.1207/s15327906mbr3901_4
- Mastrogriorgio, A. y Petracca, E. (2014). Numerals as triggers of System 1 and System 2 in the ‘bat and ball’ problem. *Mind & Society*, 13, 135-148. doi:10.1007/s11299-014-0138-8
- *Mata, A. (2016). Proportion dominance in valuing lives: The role of deliberative thinking. *Judgment and Decision Making*, 11, 441-448.
- Mata, A. y Almeida, T. (2014). Using metacognitive cues to infer others’ thinking. *Judgment & Decision Making*, 9, 349-359.
- Mata, A., Ferreira, M. B. y Sherman, S. J. (2013). The metacognitive advantage of deliberative thinkers: A dual-process perspective on overconfidence. *Journal of Personality and Social Psychology*, 105, 353-373. doi:10.1037/a0033640
- Mata, A., Fiedler, K., Ferreira, M. B. y Almeida, T. (2013). Reasoning about others' reasoning. *Journal of Experimental Social Psychology*, 49, 486-491. doi:10.1016/j.jesp.2013.01.010
- Mata, A., Schubert, A. L. y Ferreira, M. B. (2014). The role of language comprehension in reasoning: How “good-enough” representations induce biases. *Cognition*, 133, 457-463. doi:10.1016/j.cognition.2014.07.011
- Mendonça, C. S. P. (2012). *Cognitive Reflection Test e teorias dualistas no raciocínio e nas atitudes* (tesis de máster). Recuperado de: <http://hdl.handle.net/10451/7809>
- Metcalfe, J. y Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, 106, 3-19. doi:10.1037/0033-295x.106.1.3
- Metis (2001). *Manual técnico de la Bateria de Habilidades Cognitivas*. Manuscrito no publicado.
- Meub, L. y Proeger, T. (2017). The impact of communication regimes and cognitive abilities on group rationality: Experimental evidence. *Journal of Economic Behavior & Organization*, 135, 229-238. doi:10.1016/j.jebo.2017.02.010
- Meyer, A., Frederick, S., Burnham, T. C., Guevara Pinto, J. D., Boyer, T. W., Ball, L. J., ... y Schuldt, J. P. (2015). Disfluent fonts don't help people solve math problems. *Journal of Experimental Psychology: General*, 144, e16-e30. doi:10.1037/xge0000049
- Mischel, W. (1996). From good intentions to willpower. (1996). En P. Gollwitzer y J. Bargh (Eds.), *The psychology of action* (pp. 249-292). New York: Guilford.
- Morewedge, C. K. y Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, 14, 435-440. doi:10.1016/j.tics.2010.07.004
- Moritz, B. B., Hill, A. V. y Donohue, K. L. (2013). Individual differences in the newsvendor problem: Behavior and cognitive reflection. *Journal of Operations Management*, 31, 72-85. doi:10.1016/j.jom.2012.11.006

- *Moritz, B., Siemsen, E. y Kremer, M. (2014). Judgmental forecasting: Cognitive reflection and decision speed. *Production and Operations Management*, 23, 1146-1160. doi:10.1111/poms.12105
- *Morsanyi, K., Busdraghi, C. y Primi, C. (2014). Mathematical anxiety is linked to reduced cognitive reflection: A potential road from discomfort in the mathematics classroom to susceptibility to biases. *Behavioral and Brain Functions*, 10, 31. doi:10.1186/1744-9081-10-31
- *Morsanyi, K., McCormack, T. y O'Mahony, E. (2017). The link between deductive reasoning and mathematics. *Thinking & Reasoning*, 24, 234-257. doi:10.1080/13546783.2017.1384760
- *Morsanyi, K., Primi, C., Handley, S. J., Chiesi, F. y Galli, S. (2012). Are systemizing and autistic traits related to talent and interest in mathematics and engineering? Testing some of the central claims of the empathizing–systemizing theory. *British Journal of Psychology*, 103, 472-496. doi:10.1111/j.2044-8295.2011.02089.x
- Mosier, C. I. (1943). On the reliability of a weighted composite. *Psychometrika*, 8, 161-168. doi:10.1007/BF02288700
- Murgui, S. y Jiménez, T. I. (2013). Efecto de supresión y mediación en el contexto de la intervención psicosocial: Diferencias, similitudes y ejemplos. *Psychosocial Intervention*, 22, 55-59. doi:10.5093/in2013a7
- *Narayanan, A. y Moritz, B. B. (2015). Decision making and cognition in multi-echelon supply chains: An experimental study. *Production and Operations Management*, 24, 1216-1234. doi:10.1111/poms.12343
- Neisser, U. (1963). The multiplicity of thought. *British Journal of Psychology*, 54, 1-14. doi:10.1111/j.2044-8295.1963.tb00857.x
- Newell, B. R., Koehler, D. J., James, G., Rakow, T. y van Ravenzwaaij, D. (2013). Probability matching in risky choice: The interplay of feedback and strategy availability. *Memory & Cognition*, 41, 329-338. doi:10.3758/s13421-012-0268-3
- Nieuwenstein, M. y van Rijn, H. (2012). The unconscious thought advantage: Further replication failures from a search for confirmatory evidence. *Judgment & Decision Making*, 7, 779-798.
- Nisbett, R. E., Peng, K., Choi, I. y Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, 108, 291-310. doi:10.1037/0033-295X.108.2.291
- Noori, M. (2016). Cognitive reflection as a predictor of susceptibility to behavioral anomalies. *Judgment and Decision Making*, 11, 114-121.

- Norman, D. A. y Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. En R. J. Davidson y G. E. Schwartz y D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research*, (pp. 1-18). Springer, Boston, MA.
- Noussair, C. N., Tucker, S. y Xu, Y. (2016). Futures markets, cognitive ability, and mispricing in experimental asset markets. *Journal of Economic Behavior & Organization*, 130, 166-179. doi:10.1016/j.jebo.2016.07.010
- *Obrecht, N. A., Chapman, G. B. y Gelman, R. (2009). An encounter frequency account of how experience affects likelihood estimation. *Memory & Cognition*, 37, 632-643. doi:10.3758/MC.37.5.632
- Oechssler, J., Roider, A. y Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, 72, 147-152. doi:10.1016/j.jebo.2009.04.018
- Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, 11, 988-1010. doi:10.3758/BF03196730
- Pacini, R. y Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, 76, 972-987. doi:10.1037//0022-3514.76.6.972
- Pearson, K. (1904). Report on certain enteric fever inoculation statistics. *British Medical Journal*, Nov 5t, 1243-1246.
- *Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J. y Fugelsang, J. A. (2014a). Cognitive style and religiosity: The role of conflict detection. *Memory & Cognition*, 42, 1-10. doi:10.3758/s13421-013-0340-7
- Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J. y Fugelsang, J. A. (2014b). The role of analytic thinking in moral judgements and values. *Thinking & Reasoning*, 20, 188-214. doi: 10.1080/13546783.2013.865000
- *Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J. y Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making*, 10, 549-563.
- *Pennycook, G., Cheyne, J. A., Koehler, D. J. y Fugelsang, J. A. (2015). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods*, 48, 341-348. doi:10.3758/s13428-015-0576-1
- *Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J. y Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition*, 123, 335-346. doi: 10.1016/j.cognition.2012.03.003
- Pennycook, G. y Ross, R. M. (2016). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology*, 7, 9. doi:10.3389/fpsyg.2016.00009

- Peters, E., Hibbard, J., Slovic, P. y Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs*, 26, 741-748. doi:10.1377/hlthaff.26.3.741
- Peters, E. y Levin, I. P. (2008). Dissecting the risky-choice framing effect: Numeracy as an individual-difference factor in weighting risky and riskless options. *Judgment and Decision Making*, 6, 435-448.
- Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K. y Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, 17, 407-413. doi:10.1111/j.1467-9280.2006.01720.x
- Piazza, J. y Sousa, P. (2014). Religiosity, political orientation, and consequentialist moral thinking. *Social Psychological and Personality Science*, 5, 334-342. doi:10.1177/1948550613492826
- Pinillos, N. Á., Smith, N., Nair, G. S., Marchetto, P. y Mun, C. (2011). Philosophy's new challenge: Experiments and intentional action. *Mind & Language*, 26, 115-139. doi:10.1111/j.1468-0017.2010.01412.x
- Pollock, J. L. (1989). Oscar: A general theory of rationality. *Journal of Experimental & Theoretical Artificial Intelligence*, 1, 209-226. doi:10.1080/09528138908953702
- Pollock, J. L. (1991). OSCAR: A general theory of rationality. En R. Cummins y J. L. Pollock (Eds.), *Philosophy and AI: Essays at the interface* (pp. 189-213). Cambridge, MA, US: The MIT Press.
- *Ponti, G. y Carbone, E. (2009). Positional learning with noise. *Research in Economics*, 63, 225-241. doi:10.1016/j.rie.2009.09.002
- Ponti, G. y Rodriguez-Lara, I. (2015). Social preferences and cognitive reflection: Evidence from a dictator game experiment. *Frontiers in Behavioral Neuroscience*, 9, 146. doi:10.3389/fnbeh.2015.00146
- *Ponti, G., Rodriguez-Lara, I. y Di Cagno, D. (2014). *Doing it now or later with payoff externalities: Experimental evidence on social time preferences* (Working Paper No. 5). Recuperado de: <http://static.luiss.it/RePEc/pdf/cesare/1401.pdf>
- *Poore, J. C., Forlines, C. L., Miller, S. M., Regan, J. R. e Irvine, J. M. (2014). Personality, cognitive style, motivation, and aptitude predict systematic trends in analytic forecasting behavior. *Journal of Cognitive Engineering and Decision Making*, 8, 374-393. doi:10.1177/1555343414554702
- Postlethwaite, B. E. (2011). *Fluid ability, crystallized ability, and performance across multiple domains: A meta-analysis* (Doctoral dissertation). The University of Iowa, Iowa. doi:10.17077/ctd.zopi8wvs

- Prieto, G. y Delgado, A. R. (2010). Fiabilidad y validez. *Papeles del Psicólogo*, 31, 67-74.
- *Primi, C., Donati, M. A., Chiesi, F. y Morsanyi, K. (2018). Are there gender differences in cognitive reflection? Invariance and differences related to mathematics. *Thinking & Reasoning*, 24, 258-279. doi:10.1080/13546783.2017.1387606
- *Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A. y Hamilton, J. (2015). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, 29, 453-469. doi:10.1002/bdm.1883
- *Primi, C., Morsanyi, K., Donati, M. A., Galli, S. y Chiesi, F. (2017). Measuring probabilistic reasoning: The construction of a new scale applying item response theory. *Journal of Behavioral Decision Making*, 30, 933-950. doi:10.1002/bdm.2011
- Raven, J., Raven, J. C. y Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Oxford, UK: Oxford Psychologists Press Ltd.
- *Razmyar, S. y Reeve, C. L. (2013). Individual differences in religiosity as a function of cognitive ability and cognitive style. *Intelligence*, 41, 667-673. doi:10.1016/j.intell.2013.09.003
- Reber, A. S. (1993). *Implicit learning and tacit knowledge. An essay on the cognitive unconscious*. New York, US: Oxford University Press.
- Reuben, E., Sapienza, P. y Zingales, L. (2015). Procrastination and impatience. *Journal of Behavioral and Experimental Economics*, 58, 63-76. doi:10.1016/j.socsc.2015.07.005
- Reyna, V. F., Helm, R. K., Weldon, R. B., Shah, P. D., Turpin, A. G. y Govindgari, S. (2018). Brain activation covaries with reported criminal behaviors when making risky choices: A fuzzy-trace theory approach. *Journal of Experimental Psychology: General*, 147, 1094-1109. doi:10.1037/xap0000434
- Reyna, V. F. y Wilhelms, E. A. (2017). The gist of delay of gratification: Understanding and predicting problem behaviors. *Journal of Behavioral Decision Making*, 30, 610-625. doi:10.1002/bdm.1977
- Richardson, M., Abraham, C. y Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138, 353-387. doi:10.1037/a0026838
- Ring, P., Neyse, L., David-Barett, T. y Schmidt, U. (2016). Gender differences in performance predictions: Evidence from the Cognitive Reflection Test. *Frontiers in Psychology*, 7, 1680. doi:10.3389/fpsyg.2016.01680
- Roth, P. L., BeVier, C. A., Switzer III, F. S. y Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology*, 81, 548-556. doi:10.1037/0021-9010.81.5.548

- *Royzman, E. B., Landy, J. F. y Goodwin, G. P. (2014). Are good reasoners more incest-friendly? Trait cognitive reflection predicts selective moralization in a sample of American adults. *Judgment and Decision Making*, 9, 176-190.
- *Royzman, E. B., Landy, J. F. y Leeman, R. F. (2015). Are thoughtful people more utilitarian? CRT as a unique predictor of moral minimalism in the dilemmatic context. *Cognitive Science*, 39, 325-352. doi:10.1111/cogs.12136
- Ruch, W. W., Stang, S. W., McKillip, R. H. y Dye, D. A. (2001). *Employee Aptitude Survey technical manual*, 2nd ed. Glendale,CA: Psychological Services Inc.
- *Ruffle, B. J. y Tobol, Y. (2017). Clever enough to tell the truth. *Experimental Economics*, 20, 130-155. doi:10.1007/s10683-016-9479-y
- Salgado, J. F. (1997). VALCOR: A program for estimating standard error, confidence intervals and probability of corrected validity. *Behavior Research, Methods, Instruments, & Computers*, 29, 464-467. doi:10.3758/BF03200601
- Salgado, J. F. (1998). *Manual técnico del Inventario de Personalidad de Cinco Factores (IP/5F)*. Santiago de Compostela: Tórculo.
- Salgado, J. F. (2010). *Escalas de desempeño académico: CDTE, CDCE y CDAN*. Manuscrito no publicado, Departamento de Psicología Social, Básica y Metodología, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Salgado, J. F. (2014a). *Cognitive Reflection Test*. Manuscrito no publicado, Departamento de Ciencia Política y Sociología, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Salgado, J. F. (2014b). *Need for Cognition Scale*. Manuscrito no publicado, Departamento de Ciencia Política y Sociología, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Salgado, J. F. (2014c). *Reliability, construct and criterion validity of the Quasipsative Personality Inventory (QI5F/Tri)*. Manuscrito no publicado, Departamento de Psicología Social, Básica y Metodología, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Salgado, J. F. (2014d). *Test de Conductas Personales (CP5)*. Manuscrito no publicado, Departamento de Ciencia Política y Sociología, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Salgado, J. F. (2015). Estimating coefficients of equivalence and stability for job performance ratings: The importance of controlling for transient error on criterion measurement. *International Journal of Selection and Assessment*, 23, 37-44. doi:10.1111/ijsa.12093

- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F. y Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European Community. *Journal of Applied Psychology*, 88, 1068-1081. doi:10.1037/0021-9010.88.6.1068
- Salgado, J. F. y Moscoso, S. (2008). Selección de personal en la empresa y las AAPP: De la visión tradicional a la visión estratégica. *Papeles del Psicólogo*, 29, 16-24.
- Salgado, J. F. y Moscoso, S. (2019). The validity of general mental ability for five performance criteria: Hunter and Hunter (1984) revisited. *Frontiers in Psychology*, under review.
- Salgado, J. F., Otero, I. y Moscoso, S. (2019). *Cognitive reflection, general mental ability, and performance*. Submitted.
- Salthouse, T. A. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology*, 48, P29-P36. doi:10.1093/geronj/48.1.P29
- Salthouse, T. A. y Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, 27, 763-776.
- Samson, G. E., Graue, M. E., Weinstein, T. y Walberg, H. J. (1984). Academic and occupational performance: A quantitative synthesis. *American Educational Research Journal*, 21, 311-321. doi:10.3102/00028312021002311
- Schmidt, F. L. y Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, 1, 199-223. doi:10.1037/1082-989X.1.2.199
- Schmidt, F. L. y Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage.
- Schmidt, F. L. y Le, H. (2004). *Software for the Hunter-Schmidt meta-analysis methods*. Iowa City, IA: Department of Management and Organizations, University of Iowa.
- Schmidt, F. L., Le, H. e Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual-differences constructs. *Psychological Methods*, 8, 206-224. doi:10.1037/1082-989X.8.2.206
- Schneider, W. y Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66. doi:10.1037/0033-295X.84.1.1
- Schneider, M. y Shor, M. (2017). The common ratio effect in choice, pricing, and happiness tasks. *Journal of Behavioral Decision Making*, 30, 976-986. doi:10.1002/bdm.2017

- *Schulze, C. y Newell, B. R. (2015). Compete, coordinate, and cooperate: How to exploit uncertain environments with social interaction. *Journal of Experimental Psychology: General*, 144, 967-981. doi:10.1037/xgc0000096
- Schwartz, L. M., Woloshin, S., Black, W. C. y Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, 127, 966-972. doi:10.7326/0003-4819-127-11-199712010-00003
- *Sevenants, A., Dicussaert, K. y Schaeken, W. (2011). Truth table tasks: Irrelevance and cognitive ability. *Thinking & Reasoning*, 17, 213-246. doi:10.1080/13546783.2011.569153
- *Shenhav, A., Rand, D. G. y Greene, J. D. (2011). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, 141, 423-428. doi:10.1037/a0025391
- *Shenhav, A., Rand, D. G. y Greene, J. (2017). The relationship between intertemporal choice and following the path of least resistance across choices, preferences, and beliefs. *Judgment and Decision Making*, 12, 1-18. doi:10.2139/ssrn.2724547
- Shepard, R. N. y Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703. doi:10.1126/science.171.3972.701
- Shiffrin, R. M. y Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127-190. doi:10.1037/0033-295X.84.2.127
- Shipley, W. C. (1986). *Shipley Institute of Living Scale*. Los Angeles, CA: Western Psychological Services.
- Shtulman, A. y McCallum, K. (2014). Cognitive reflection predicts science understanding. En P. Bello, M. Guarini, M. McShane y B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the Cognitive Science Society* (pp. 2937-2942). Austin, TX: Cognitive Science Society.
- Siegler, R. S. y Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, 14, 237-243. doi:10.1111/1467-9280.02438
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 99-118. doi:10.2307/1884852
- Simonovic, B., Stuppel, E. J., Gale, M. y Sheffield, D. (2017). Stress and risky decision making: Cognitive reflection, emotional learning or both. *Journal of Behavioral Decision Making*, 30, 658-665. doi:10.1002/bdm.1980

- Simonson, I. y Sela, A. (2011). On the heritability of consumer decision making: An exploratory approach for studying genetic effects on judgment and choice. *Journal of Consumer Research*, 37, 951-966. doi:10.1086/657022
- *Sinayev, A. y Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology*, 6, 532. doi:10.3389/fpsyg.2015.00532
- *Sirota, M. y Juanchich, M. (2011). Role of numeracy and cognitive reflection in bayesian reasoning with natural frequencies. *Studia Psychologica*, 53, 151-161.
- Sirota, M., Juanchich, M., Kostopoulou, O. y Hanak, R. (2014). Decisive evidence on a smaller-than-you-think phenomenon: Revisiting the “1-in-X” effect on subjective medical probabilities. *Medical Decision Making*, 34, 419-429. doi:10.1177/0272989X13514776
- Sjöberg, A., Sjöberg, S. y Forssén, K. (2006). *Predicting job performance. Manual*. Assessio International, Stockholm.
- *Skagerlund, K., Lind, T., Strömbäck, C., Tinghög, G. y Västfjäll, D. (2018). Financial literacy and the role of numeracy—how individuals’ attitude and affinity with numbers influence financial literacy. *Journal of Behavioral and Experimental Economics*, 74, 18-25. doi:10.1016/j.socec.2018.03.004
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22. doi:10.1037/0033-2909.119.1.3
- Smith, E. R. y DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4, 108-131. doi:10.1207/S15327957PSPR0402_01
- Society for Industrial and Organizational Psychology (2018). *Principles for the validation and use of personnel selection procedures* (5th ed.). Bowling Green, OH: Author.
- Spadaccini, J. y Esteves, J. E. (2014). Intuition, analysis and reflection: An experimental study into the decision-making processes and thinking dispositions of osteopathy students. *International Journal of Osteopathic Medicine*, 17, 263-271. doi:10.1016/j.ijosm.2014.04.004
- Spunt, R. P. (2015). Dual-process theories in social cognitive neuroscience. En A. W. Toga, (Ed.), *Brain mapping: An encyclopedic reference*, vol. 3, pp. 211-215. Academic Press: Elsevier. doi:10.1016/B978-0-12-397025-1.00181-0
- Stankov, L. (1997). *Gf-Gc Quickie Test Battery*. Sydney, NSW: Entelligence Testing Products.
- Stankov, L. (2000). Complexity, metacognition, and fluid intelligence. *Intelligence*, 28, 121-143. doi:10.1016/S0160-2896(99)00033-1
- Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.

- Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory. En J. S. B. Evans y K. Frankish (Eds.), *In two minds: Dual processes and beyond*, (pp. 55-88). New York, UU.EE: Oxford University Press.
- Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York, UU. EE: Oxford University Press.
- Stanovich, K. E. y West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127, 161-188. doi:10.1037/0096-3445.127.2.161
- Stanovich, K. E. y West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645-665. doi:10.1017/S0140525X00003435
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Viking.
- Strack, F. y Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220-247. doi:10.1207/s15327957pspr0803_1
- Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, 35, 401-426. doi:10.1016/j.intell.2006.09.004
- *Stupple, E. J. N., Gale, M. y Richmond, C. (2013). Working memory, cognitive miserliness and logic as predictors of performance on the Cognitive Reflection Test. En M. Knauff, M. Pauen, N. Sebanz y I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the Cognitive Science Society*: Vol. 35, No. 35. Cognitive Science Society: Austin, TX. Recuperado de: https://www.academia.edu/3656911/Working_Memory_Cognitive_Miserliness_and_Logic_as_Predictors_of_Performance_on_the_Cognitive_Reflection_Test
- Sutherland, S. L. y Cimpian, A. (2015). An explanatory heuristic gives rise to the belief that words are well suited for their referents. *Cognition*, 143, 228-240. doi:10.1016/j.cognition.2015.07.002
- Svedholm-Häkkinen, A. M. y Lindeman, M. (2018). Actively open-minded thinking: Development of a shortened scale and disentangling attitudes towards knowledge and people. *Thinking & Reasoning*, 24, 21-40. doi:10.1080/13546783.2017.1378723
- *Szász, B., Szollosi, A., Palfi, B. y Aczel, B. (2017). The cognitive reflection test revisited: Exploring the ways individuals solve the test. *Thinking & Reasoning*, 23, 207-234. doi:10.1080/13546783.2017.1292954
- Szollosi, A., Bago, B., Szász, B. y Aczel, B. (2017). Exploring the determinants of confidence in the bat-and-ball problem. *Acta Psychologica*, 180, 1-7. doi:10.1016/j.actpsy.2017.08.003

- Tangney, J. P., Baumeister, R. F. y Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72, 271-324. doi:10.1111/j.0022-3506.2004.00263.x
- Taylor, M. P. (2016). Are high-ability individuals really more tolerant of risk? A test of the relationship between risk aversion and cognitive ability. *Journal of Behavioral and Experimental Economics*, 63, 136-147. doi:10.1016/j.soccec.2016.06.001
- *Teovanović, P., Knežević, G. y Stankov, L. (2015). Individual differences in cognitive biases: Evidence against one-factor theory of rationality. *Intelligence*, 50, 75-86. doi:10.1016/j.intell.2015.02.008
- Thaler, R. H. y Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy*, 89, 392-406. doi:10.1086/260971
- The Psychological Corporation Ltd. (1974). *General Clerical Test Battery (GCT)*. Adaptación y traducción española, Madrid: TEA Ediciones.
- Thompson, V. A. (2009). Dual process theories: A metacognitive perspective. En J. S. B. Evans y K. Frankish (Eds.), *In two minds: Dual processes and beyond*, (pp. 171-195). New York, UU.EE: Oxford University Press.
- Thompson, V. A. (2010). Towards a dual-process model of conditional inference. En M. Oaksford (Ed.), *The psychology of conditionals*. Oxford: Oxford University Press.
- Thompson, V. A., Pennycook, G., Trippas, D. y Evans, J. S. B. (2018). Do smart people have better intuitions? *Journal of Experimental Psychology: General*, 147, 945-961. doi:10.1037/xge0000457
- Thompson, V. A., Turner, J. A. P. y Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, 63, 107-140. doi:10.1016/j.cogpsych.2011.06.001
- *Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y. y Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, 128, 237-251. doi:10.1016/j.cognition.2012.09.012
- *Thomson, K. S. y Oppenheimer, D. M. (2016). Investigating an alternate form of the Cognitive Reflection Test. *Judgment and Decision Making*, 11, 99-113.
- Thorndike, R. L. (1949). *Personnel selection. Test and measurement techniques*. New York: John Wiley & Sons, Inc.
- Timmerman, M. E. y Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16, 209-220. doi:10.1037/a0023353

- Toates, F. (2006). A model of the hierarchy of behaviour, cognition, and consciousness. *Consciousness and Cognition*, 15, 75-118. doi:10.1016/j.concog.2005.04.008
- *Toplak, M. E., West, R. F. y Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39, 1275-1289. doi:10.3758/s13421-011-0104-1
- *Toplak, M. E., West, R. F. y Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20, 147-168. doi:10.1080/13546783.2013.844729
- *Toplak, M. E., West, R. F. y Stanovich, K. E. (2017). Real-world correlates of performance on heuristics and biases tasks in a community sample. *Journal of Behavioral Decision Making*, 30, 541-554. doi:10.1002/bdm.1973
- Travers, E., Rolison, J. J. y Feeney, A. (2016). The time course of conflict on the Cognitive Reflection Test. *Cognition*, 150, 109-118. doi:10.1016/j.cognition.2016.01.015
- Trémolière, B., De Neys, W. y Bonnefon, J. F. (2014). The grim reasoner: Analytical reasoning under mortality salience. *Thinking & Reasoning*, 20, 333-351. doi:10.1080/13546783.2013.823888
- *Trippas, D., Pennycook, G., Verde, M. F. y Handley, S. J. (2015). Better but still biased: Analytic cognitive style and belief bias. *Thinking & Reasoning*, 21, 431-445. doi:10.1080/13546783.2015.1016450
- Trouche, E., Sander, E. y Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, 143, 1958-1971. doi:10.1037/a0037099
- Turner, M. L. y Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154. doi:10.1016/0749-596X(89)90040-5
- Tversky, A. y Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131. doi:10.1126/science.185.4157.1124
- Unsworth, N., Heitz, R. P., Schrock, J. C. y Engle, R. W. (2005). An automated version of the Operation Span Task. *Behavior Research Methods*, 37, 498-505. doi:10.3758/BF03192720
- van der Heijden, E., Klein, T. J., Müller, W. y Potters, J. (2012). Framing effects and impatience: Evidence from a large scale experiment. *Journal of Economic Behavior & Organization*, 84, 701-711. doi:10.1016/j.jebo.2012.09.017
- van Prooijen, J. W. (2017). Why education predicts decreased belief in conspiracy theories. *Applied Cognitive Psychology*, 31, 50-58. doi:10.1002/acp.3301

- Vandenberg, S. G. y Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599–604. doi:10.2466/pms.1978.47.2.599.
- *Ventis, L. (2015). Thinking fast and slow in the experience of humor. *Humor*, 28, 351-373. doi:10.1515/humor-2015-0070
- Viator, R. E., Bagley, P. L., Barnes, B. G. y Harp, N. L. (2014). Measuring reflective cognitive capacity: A methodological recommendation for accounting research of feedback effects. *Behavioral Research in Accounting*, 26, 131-160. doi:10.2308/bria-50803
- Wason, P. C. (1966). Reasoning. En, B. M. Foss (Ed.), *New horizons in psychology* (pp. 135-151). Harmonds-worth: Penguin.
- Wason, P. C. y Evans, J. S. B. (1975). Dual processes in reasoning? *Cognition*, 3, 141-154. doi:10.1016/0010-0277(74)90017-1
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX.: The Psychological Corporation.
- Wechsler, D. (1987). *Wechsler Adult Intelligence Scale – III*. New York: The psychological Corporation.
- Weinhardt, J. M., Hendijani, R., Harman, J. L., Steel, P. y Gonzalez, C. (2015). How analytic reasoning style and global thinking relate to understanding stocks and flows. *Journal of Operations Management*, 39, 23-30. doi:10.1016/j.jom.2015.07.003
- *Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J. y Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making*, 26, 198-212. doi:10.1002/bdm.1751
- *Welsh, M., Burns, N. y Delfabbro, P. (2013). The Cognitive Reflection Test: how much more than numerical ability? En *Proceedings of the annual meeting of the Cognitive Science Society*: Vol. 35 (pp. 1587-1592). Recuperado de: <https://cloudfront.escholarship.org/dist/prd/content/qt68n012fh/qt68n012fh.pdf>
- *West, R. F., Meserve, R. J. y Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of Personality and Social Psychology*, 103, 506-519. doi:10.1037/a0028857
- Wherry, R. (1931). A new formula for predicting the shrinkage of the coefficient of multiple correlation. *The Annals of Mathematical Statistics*, 2, 440-457. doi:10.1214/aoms/1177732951
- *Willard, A. K. y Norenzayan, A. (2017). “Spiritual but not religious”: Cognition, schizotypy, and conversion in alternative beliefs. *Cognition*, 165, 137-146. doi:10.1016/j.cognition.2017.05.018

- Wilson, T. D., Lindsey, S. y Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126. doi:10.1037//0033-295X.107.1.101
- Witkin, H. A., Oltman, P. K., Raskin, E. y Karp, S. A. (1971). *A manual for the Embedded Figures Test*. Palo Alto, CA: Consulting Psychologists Press.
- Wolf, M. B. y Ackerman, P. L. (2005). Extraversion and intelligence: A meta-analytic investigation. *Personality and Individual Differences*, 39, 531-542. doi:0.1016/j.paid.2005.02.020
- Wolf, B., Momirović, K. y Džamonja, Z. (1992). *Intelligence Test Battery—KOG 3*. Belgrade: Center for Applied Psychology.
- Wonderlic y Asociados (1998). *Wonderlic Personnel Test and Scholastic Level Exam. User's Manual*. Libertyville, IL: Wonderlic, Inc
- Woodcock, R.W., McGrew, K. S. y Mather, N. (2001). *Woodcock-Johnson Tests of Achievement*. Itasca, IL: Riverside Publishing.
- *Yılmaz, O. y Sarıbay, S. A. (2016). An attempt to clarify the link between cognitive style and political ideology: A non-western replication and extension. *Judgment and Decision Making*, 11, 287-300. doi:10.1017/CBO9781107415324.004
- *Yılmaz, O. y Sarıbay, S. A. (2017). The relationship between cognitive style and political orientation depends on the measures used. *Judgment and Decision Making*, 12, 140-147.
- Yin, P. y Fan, X. (2001). Estimating R^2 shrinkage in multiple regression: A comparison of different analytical methods. *The Journal of Experimental Education*, 69, 203-224. doi:10.1080/00220970109600656
- York, T. T., Gibson, C. y Rankin, S. (2015). Defining and Measuring Academic Success. *Practical Assessment, Research & Evaluation*, 20, 1-36.

APÉNDICES





APÉNDICE A.

Distribución e Integración de las Puntuaciones del TRC Original

Tabla A1.
Distribución de las Puntuaciones del TRC Original

Autores (año)	Media	SD	N	r_{xx}	Observaciones
Frederick (2005)	1.24	1.09	3,428	-	11 muestras independientes
*Guthrie, Rachlinski y Wistrich (2007)	1.23	1.04	252	-	
*Oechssler, Roeder y Schmitz (2009)	2.05	1.00	564	-	
*Campitelli y Labollita (2010)	0.66	0.90	157	-	
Del Missier, Mântyla y Bruine de Bruin (2011)	0.67	0.92	212	.55	
*Hoppe y Kusterer (2011)	1.84	1.05	411	-	
Pinillos, Smith, Näir, Marchetto y Mun (2011)	0.77	1.00	463		Completaron el TRC antes
Sevenants, Dieussaert y Schaeken (2011)	0.85	1.10	405		Completaron el TRC después
	1.75	1.08	91	-	Experimento 1
	1.38	1.12	77	-	Experimento 2
*Simonson y Sela (2011)	0.66	-	220	-	Gemelos monocigóticos
	0.78	-	140	-	Gemelos dicigóticos
*Sirota y Juanchich (2011)	1.30	1.10	94	.60	
*Toplak, West y Stanovich (2011)	0.70	0.93	346	-	

Continúa

Tabla A1.
Continuación

Autores (año)	Media	SD	N	r_{xx}	Observaciones
Böckenholt (2012)	2.20	-	132	-	Hombres por la mañana
	1.70	-	143	-	Hombres por la noche
	1.80	-	148	-	Mujeres por la mañana
	1.40	-	156	-	Mujeres por la noche
*Brañas-Garza, García-Muñoz y Hernán González (2012)	0.44	0.70	191	-	
Cesarini, Johannesson, Magnusson y Wallace (2012)	0.80	1.13	5,323	-	
Gino y Ariely (2012)	1.22	1.10	97	-	Experimento 1
Iyer, Koleva, Graham, Ditto y Haidt (2012)	1.73	-	7,384	-	Liberales
	1.57	-	1,267	-	Conservadores
	2.06	-	1,070	-	Libertarios
*Liberali, Reyna, Furlan, Stein y Pardo (2012)	1.10	1.17	259	.74	Estudio 1
	1.50	1.12	190	.64	Estudio 2
Morsanyi, Primi, Handley, Chiesi y Galli (2012)	1.31	1.10	55	-	Hombres
	0.71	0.92	70	-	Mujeres
*Nieuwenstein y van Rijn (2012)	1.65	-	40	-	Hombres
	1.14	-	88	-	Mujeres
West, Meserve y Stanovich (2012)	0.79	0.92	482	-	Estudio 1
	1.48	1.25	265	-	Estudio 2
Baldi, Iannello, Riva y Antonietti (2013)	1.29	1.15	167	.68	
*Carpenter, Graham y Wolf (2013)	1.79	1.05	422	-	Experimento 1

Continúa

Tabla A1.
Continuación

Autores (año)	Media	SD	N	r_{xx}	Observaciones
Fernbach, Sloman, Louis y Shube (2013)	1.03	1.03	167	.80	Experimento 1
	1.19	1.04	201	.69	Experimento 2
*Fosgaard, Hansen y Piovesan (2013)	1.26	-	65	-	Mujeres
	1.39	-	144	-	Hombres
*Goodman, Cryder y Cheema (2013)	1.17	-	107	-	Estudio 1. Muestra <i>Mechanical Turk</i>
	0.96	-	60	-	Estudio 1. Muestra comunitaria
	1.23	-	207	-	Estudio 2. Muestra <i>Mechanical Turk</i>
	1.70	-	131	-	Estudio 2. Muestra E. universitarios
*Kahan (2013)	0.65	0.95	1,750	-	
*Lubian y Untertrefler (2013)	1.64	1.14	376	-	
Mata, Ferreira y Sherman (2013)	1.16	1.14	86	-	Estudio 1
	1.13	1.11	75	-	Estudio 2
	1.16	1.19	142	-	Estudio 6A
	1.72	1.24	80	-	Estudio 6B
*Moritz, Hill y Donohue (2013)	1.50	1.13	313	-	Estudio 1
Raznyar y Reeve (2013)	0.59	0.91	150	-	
Stupples, Gale y Richmond (2013)	1.32	1.11	65	-	Experimento 1
	1.12	1.14	49	-	Experimento 2
Weller et al. (2013)	0.83	0.99	1,970	.60	Estudio 1

Continúa

Tabla A1.
Continuación

Autores (año)	Media	SD	N	r_{xx}	Observaciones
Welsh, Burns y Delfabbro (2013)	0.73	0.94	22	-	No universitarios
	1.24	1.05	58	-	Estudiantes universitarios
	1.50	1.06	22	-	Universitarios postgraduados
Albaity, Rahman y Shahidul (2014)	0.88	-	880	-	
*Bosch-Doménech, Brañas-Garza y Espín (2014)	0.96	1.03	260	-	Muestra hombres
	0.58	0.86	363	-	Muestra mujeres
Brosnan, Hollinworth, Antoniadou y Lewton (2014)	1.42	1.10	25	-	Muestra hombres
	0.83	1.08	43	-	Muestra mujeres
*Browne, Pennycook, Goodwin y McHenry (2014)	0.46	-	1,137	-	
*Campitelli y Gerrans (2014)	0.94	1.06	2,019	.66	
*Coates y Blaszczyński (2014)	2.16	0.83	19	-	"Quitters"
	1.14	1.22	29	-	"Non-quitters"
Costa, Foucart, Arnon, Aparici y Apesteeguia (2014)	0.61	0.84	153	-	Condición nativa (español/inglés)
	0.77	0.98	151	-	Condición extranjera (español/inglés)
	1.02	1.04	159	-	Condición nativa (inglés/español)
	1.03	1.07	167	-	Condición extranjera (inglés/español)
*Guillen y Hakimov (2014)	1.28	-	180	-	Condición nativa (inglés/español)
Morsanyi, Busdraghi y Primi (2014)	0.90	0.94	328	.57	Condición nativa (inglés/español)
	0.64	0.96	184	.68	Condición extranjera (inglés/español)

Continúa

Tabla A1.
Continuación

Autores (año)	Media	SD	N	<i>r</i>_{xx}	Observaciones
*Pennycook, Cheyne, Barr, Kochler y Fugelsang (2014a)	0.36	0.37	200	-	
Poore, Forlines, Miller, Regan e Irvine (2014)	1.62	1.16	762	-	
*Royzman, Landy y Goodwin (2014)	1.34	1.20	525	.75	Estudio 1
*Sirota, Juanchich, Kostopoulou y Hanak (2014)	0.80	1.00	246	.60	Experimento 2
Spadaccini y Esteves (2014)	-	-	76	.83	
*Toplak, West y Stanovich (2014)	0.49	0.85	160	-	
Viator, Bagley, Barnes y Harp (2014)	1.02	1.07	44	-	Experimento 3
	1.30	1.09	117	-	Experimento 4
Aczel, Bago, Szollosi, Foldes y Lukacs (2015)	1.37	1.11	474	.63	Estudio 1. Grupo 1
	1.27	1.08	390	.61	Estudio 1. Grupo 2
	1.31	1.18	178	.72	Estudio 2. Grupo 1
	1.27	1.13	176	.64	Estudio 2. Grupo 2
	1.40	1.21	177	.74	Estudio 2. Grupo 3
Baghestanian, Lugovskyy y Puzzello (2015)	1.23	1.11	44	-	
Baron, Scott, Fincher y Metz (2015)	1.66	-	103	.62	Estudio 1
	1.31	-	82	-	Estudio 2
	1.48	-	104	-	Estudio 3
	1.37	-	101	-	Estudio 5

Continúa

Tabla A1.
Continuación

Autores (año)	Media	SD	N	<i>r</i>_{xx}	Observaciones
*Brañas-Garza, Kujal y Lenkei (2015)	1.20	1.13	44,558	-	^j Meta-análisis de 118 estudios
Calvillo y Burgeno (2015)	-	-	306	.63	Estudio 2
*Corgnet, Espin, Hernán-González, Kujal y Rässenti (2015)	1.95	-	38	-	Estudio 1. Muestra hombres
Corgnet, Hernán-González y Mateo (2015)	1.52	-	42	-	Estudio 1. Muestra mujeres
*Duttile (2015)	1.22	1.10	264	-	
*Duttile e Inukai (2015)	1.62	-	100	-	
	1.58	1.14	66	-	Estudio 1
	1.98	1.03	44	-	Estudio 2. Muestra alemanes
	2.48	0.77	44	-	Estudio 2. Muestra japoneses
*Fehr y Huck (2015)	1.49	-	240	-	
Gervais (2015)	0.72	1.05	757	-	Estudio 1
	0.78	1.06	567	-	Estudio 2
Grafteo, Polonio y Bonini (2015)	1.40	1.31	35	-	Experimento 2
*Insler, Compton y Schmitt (2015)	1.78	1.12	364	-	
^b Meyer et al. (2015)	1.43	-	3,657	-	Fuente fluida
	1.42	-	3,710	-	^j Fuente no fluida
*Narayanan y Moritz (2015)	0.97	1.04	908	-	
*Pennycook, Cheyne, Kochler y Fugelsang (2015)	1.04	1.05	372	-	
Ponti y Rodríguez-Lara (2015)	1.28	1.18	96	-	

Continúa

Tabla A1.

Continuación

Autores (año)	Media	SD	N	<i>r</i>_{xx}	Observaciones
Primi, Morsanyi, Chiesi, Donati y Hamilton (2015)	1.24	-	438	.65	^J Estudio 1. Límite de tiempo
	1.37	-	988	-	Estudio 2
	0.61	-	70	-	Estudio 3. Muestra adolescentes
	1.29	-	287	-	Estudio 3. Muestra adultos
Sutherland y Cimpian (2015)	1.66	1.13	128	-	Estudio 5
Teovanović, Knežević y Stankov (2015)	0.20	0.50	243	.39	
Ventis (2015)	1.70	1.10	148	.67	
Weinhardt, Hendijani, Harman, Steel y Gonzalez (2015)	1.29	1.11	1,045	-	
Alós-Ferrer, Garagnani y Hügelschäfer (2016)	2.03	1.01	155	-	
*Alós-Ferrer y Hügelschäfer (2016)	1.13	1.15	416	-	Experimento 1
	1.30	1.09	111	-	Experimento 2
	1.98	1.09	364	-	Experimento clase
Eriksson y Jansson (2016)	1.52	1.17	180	-	Estudio 1
*Hanaki, Jacquemet, Luchini y Zylbersztejn (2016)	0.48	0.85	200	-	
Juanchich, Dewberry, Sirota y Narendran (2016)	0.97	1.11	384	.71	
Kamei (2016)	1.40	-	300	-	
*Kiss, Rodriguez-Lara y Rosa-García (2016)	0.57	0.98	60	-	
Koscielniak, Rydzewska y Sedek (2016)	1.00	1.04	81	-	Jóvenes-adultos (18-23)
	0.58	0.70	77	-	Adultos-mayores (65-80)
Kramer (2016)	1.50	-	467	-	

Continúa

Tabla A1.
Continuación

Autores (año)	Media	SD	N	r_{xx}	Observaciones
Lindeman y Svedholm-Häkkinen (2016)	1.87	1.09	258	-	
*Lohse (2016)	1.81	1.08	284	-	
Noori (2016)	1.89	1.12	395	-	
Noussair, Tucker y Xu (2016)	1.51	-	216	-	^j 3 min. para responder
Ring, Neyse, David-Barett y Schmidt (2016)	1.76	1.07	131	-	^j 10 min. para responder
Taylor (2016)	1.15	-	48	-	Tratamiento 1
	1.00	-	45	-	Tratamiento 2
	1.04	-	46	-	Tratamiento 3
	1.19	-	42	-	Tratamiento 4
Thomson y Oppenheimer (2016)	-	-	134	.62	Estudio 2
	1.90	-	144	-	^j Prevía exposición al TRC
	1.29	-	56	-	No previa exposición al TRC
Yilmaz y Sarıbay (2016)	0.91	1.11	337	.72	Estudio 1
	1.96	1.00	691	.50	Estudio 2
	0.93	1.18	126	.43	Estudio 3A
	0.94	1.09	86	.70	Estudio 3B
Agnew y Harrison (2017)	1.82	-	425	-	
Aktaş, Yılmaz y Bağcıkapılı (2017)	1.30	1.00	269	.54	Estudio 1
	1.36	1.07	246	.63	Estudio 2
	0.80	1.04	389	.73	Estudio 3

Continúa

Tabla A1.
Continuación

Autores (año)	Media	SD	N	<i>r</i>_{xx}	Observaciones
Drummond y Fischhoff (2017)	1.80	1.20	397	.80	Estudio 1a
	1.70	1.20	395	.73	Estudio 1c
	1.50	1.30	270	.80	Estudio 2
Meub y Proeger (2017)	1.65	1.14	131	-	
Primi, Morsanyi, Donati, Galli y Chiesi (2017)	1.59	1.16	560	-	
Reyna y Wilhelms (2017)	1.42	1.12	845	.64	Estudio 2
	1.53	1.11	393	.60	Estudio 3
Ruffle y Tobol (2017)	1.33	1.04	427	-	
Schneider y Shor (2017)	1.55	1.16	158	-	
Shenhav, Rand y Greene (2017)	1.18	1.15	2,329	.70	
Szaszi, Szollosi, Palfi y Aczel (2017)	0.80	1.00	210	.64	
van Prooijen (2017)	0.96	1.01	970	-	Estudio 2
Yilmaz y Saribay (2017)	-	-	426	.75	
Königsheim, Lukas y Nöth (2018)	1.98	1.04	98	-	
Reyna, Helm, Weldon, Shah, Turpin y Govindgari (2018)	1.45	1.09	32	-	
Skagerlund, Lind, Strömbäck, Tinghög y Västfjäll (2018)	0.90	1.02	2,063	.63	
Estudio 3 de esta tesis	1.42	1.08	1,367	.71 ^c	

Nota. SD = desviación típica; N = tamaño de la muestra de cada estudio singular; r_{xx} = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach del TRC.

*Los estudios marcados con asterisco son algunos de los estudios singulares integrados en el meta-análisis de Brañas-Garza, Kujal y Lenkei (2015). Estos estudios no se incluyen en el MA_2 .

^a Este MA no se incluye en el MA_1 ni en el MA_3 .

^b Para poder incluir este estudio en el MA_2 , se han recalculado los datos eliminando la muestra de Alter et al. (2007), por estar ésta ya integrada en el MA de Brañas-Garza et al. (2015). La media obtenida para la muestra de fluidez perceptual fue de 1.43 ($N = 3,637$) y para la muestra de no fluidez 1.41 ($N = 3,690$).

^c Coeficiente de fiabilidad test-retest. No integrado en el meta-análisis.

^j Estos estudios aportan fuentes de error no comunes y no fueron integrados en el MA_3 .

Tabla A2.
Integración de las Puntuaciones del TRC Original

	$K_{muestras}$	$K_{estudios}$	N	\bar{X}_w	$S^2_{\bar{X}}$	SD_{ac}	S^2_{SD}	\bar{r}_{xx}
MA_1	157	125	72,144	1.25	.16	1.15	.011	.66
MA_2	138	115	101,314	1.25	.08	1.14	.004	.66
MA_3	148	121	66,572	1.22	.17	1.15	.010	.66

Nota. $K_{muestras}$ = número de muestras independientes integradas; $K_{estudios}$ = número de estudios integrados; N = tamaño de la muestra total; \bar{X}_w = puntuación promedio ponderada por el tamaño de la muestra; $S^2_{\bar{X}}$ = varianza observada en el conjunto de puntuaciones ponderadas por el tamaño de la muestra; SD_{ac} = desviación típica acumulada o desviación típica de la población; S^2_{SD} = varianza de las desviaciones típicas de las puntuaciones; r_{xx} = coeficiente de fiabilidad alfa promedio; MA_1 = integra todos los estudios individuales, incluso los que incluyen fuentes de error no comunes. No incluye el MA de Brañas-Garza et al. (2015); MA_2 = integra todos los estudios individuales incluso, los que incluyen fuentes de error no comunes. Incluye el MA de Brañas-Garza et al. (2015), por lo que no incluye los estudios individuales integrados en ese MA ; MA_3 = integra todos los estudios individuales, excepto los que incluyen errores no comunes. No incluye el MA de Brañas-Garza et al. (2015).

Aclaraciones sobre el Procedimiento Seguido en el Cálculo del Meta-análisis

Se han llevado a cabo 3 meta-análisis en función de las características de los estudios localizados (véase Tabla A2). Los estudios primarios han sido clasificados en dos tipos: (a) estudios que incluyen únicamente fuentes de error comunes, es decir, los resultados están afectados por errores debidos a la idiosincrasia de los participantes, errores debido al estado de ánimo de los sujetos, errores debidos a la estructura interna del propio test y errores aleatorios o (b) estudios que incluyen, además, fuentes de error no comunes, es decir, aquellos en los que existen factores externos que pueden afectar al desempeño en el TRC. Ejemplos de fuentes de error no comunes serían establecer un tiempo límite de respuesta al test (Guillen y Hakimov, 2014; Noussair et al., 2016; Primi et al., 2015), manipular la fluidez de la fuente (Meyer et al., 2015), el idioma de presentación de los ítems (Costa, Foucart, Arnon et al., 2014) o haber estado previamente expuesto al test (Thomson y Oppenheimer, 2016). De este modo, el primer meta-análisis que presenta la Tabla A2 integra los estudios singulares con fuentes de error comunes y fuentes de error no comunes pero no incluye el MA de Brañas-Garza et al. (2015), para controlar los efectos que este estudio pueda tener sobre los resultados. El segundo meta-análisis incluye estudios con fuentes de error comunes y no comunes y el meta-análisis de Brañas-Garza et al. (2015). En este caso se eliminaron, de entre los estudios primarios integrados, los estudios incluidos en el meta-análisis de Brañas-Garza y colaboradores para evitar la duplicidad de datos. Del mismo modo, el estudio de Meyer et al. (2015) fue recalculado eliminando el estudio de Alter et al. (2007) integrado también en el meta-análisis de Brañas-Garza y colaboradores. Finalmente, la tercera integración incluye únicamente los estudios individuales que aportan fuentes de error comunes. Los resultados de este meta-análisis son los más ajustados ya que controla los resultados por más fuentes de error.



APÉNDICE B.

Nuevos Ítems de RC Desarrollados/Adaptados por otros Autores

Finucane y Gullion (2010):

- Si 2 enfermeras necesitan 2 minutos para medir la presión arterial de 2 pacientes, ¿cuánto tiempo necesitarían 200 enfermeras para medir la presión arterial de 200 pacientes?
- Un plato de sopa y de ensalada cuestan 5.50 dólares en total. El plato de sopa cuesta un dólar más que la ensalada. ¿Cuánto cuesta la ensalada?
- Sally está preparando té. Cada hora, la concentración del té se duplica. Si se requiere 6 horas para que el té esté listo, ¿cuánto tiempo sería necesario para que el té alcance la mitad de la concentración final?

Böckenholt (2012):

- Jack está mirando a Anne, pero Anne está mirando a George. Jack está casado, pero George no. ¿Está una persona casada mirando a una persona soltera? (A) sí (B) no (C) no se puede determinar.
Nota. Desarrollado por Stanovich, 2009. No implica capacidad numérica.
- Un ciclista completa un recorrido de diez millas a una velocidad media de 10 millas/hora. ¿Qué tan rápido debe circular el ciclista en la segunda ronda para alcanzar una velocidad media de 20 millas/hora en ambas rondas?

Mendonça (2012):

- Si fuera ciego, sordo y sin voz, ¿cuántos sentidos todavía poseería?
- Divida 30 por $\frac{1}{2}$ y agregue 20 al resultado. ¿Cuál es el valor final?
- 800 gallinas producen 800 huevos en ocho días. ¿Cuántos huevos producen 400 gallinas en 4 días?
- Dos jugadores A y B tienen cada uno siete monedas. Si el jugador A pierde una moneda a favor del jugador B, ¿con cuántas monedas más se queda el jugador B?

Kinnunen y Windmann (2013):

- ¿Cuánto pesaría un ladrillo si este pesa 1Kg y medio ladrillo?

Mata, Fiedler, Ferreira y Almeida (2013):

- Un televisor y un DVD juntos cuestan 110 dólares. La televisión cuesta 100 dólares más que el DVD. ¿Cuánto cuesta el DVD?
Nota. Desarrollaron dos ítems adicionales no recogidos expresamente en el artículo (son adaptaciones de los ítems de la máquina y del lago de Frederick, 2005).

Borghans y Golsteyn (2014):

- Juntos, una pelota y una gorra cuestan 1.10 euros. La pelota cuesta 1.00 euro más que la gorra. ¿Cuánto cuesta la gorra?
- Si lanzas una moneda al aire dos veces, ¿qué probabilidad hay de que aparezca "cara" al menos una vez?
- Dos coches que circulan por el mismo carril se aproximan entre sí. El coche A conduce a una velocidad de 120 km/h., el coche B a 60 km/h. ¿Qué distancia habrá entre los dos coches 1 minuto antes de que choquen?

- Si Timo bebe una botella de agua en 6 días y Esther tarda 12 días en terminar la misma botella, ¿cuánto tiempo pasará antes de que terminen una botella juntos?
- Si tres vendedores envuelven seis juguetes en media hora, ¿cuántos vendedores se necesitarían para envolver 20 juguetes en 1 hora?
- En una carrera, Bart se posiciona en el puesto 15º y, simultáneamente, en el último 15º puesto. ¿Cuántas personas participaron en la carrera?

Grossman, van der Weele y Andrijevnik (2014):

- Una raqueta y una pelota cuestan en total 1.10 euros. La raqueta cuesta 1.00 euro más que la pelota. ¿Cuánto cuesta la pelota?
- Un médico le da tres píldoras y le dice que tome una cada media hora. ¿Cuántos minutos tarda en no tener ninguna píldora?

Mata y Almeida (2014):

- Un televisor y un DVD juntos cuestan 88 dólares. La televisión cuesta 80 dólares más que el DVD. ¿Cuánto cuesta el DVD?
- Si 10 gallinas tardan 10 días en poner 10 huevos, ¿cuánto tiempo tardarán 100 gallinas en poner 100 huevos?
- Un virus informático está infectando el sistema de un ordenador. Cada minuto, la cantidad de archivos infectados se duplica. Si el virus tarda 100 minutos en infectar todo el sistema, ¿cuánto tiempo tardará en infectar la mitad del sistema?
- Divida 40 por $\frac{1}{2}$ y agregue 5. ¿Cuál es el resultado final?

Shtulman y McCallum (2014):

- Una casa tiene un salón y un estudio que son perfectamente cuadrados. El salón tiene 4 veces los metros cuadrados del estudio. Si las paredes del estudio miden 10 pies de largo, ¿Qué largo tienen las paredes de la sala de estar?
- El dueño de una tienda redujo en un 10% el precio de un par de zapatos de 100 dólares. Una semana después, redujo el precio de los zapatos en otro 10%. ¿Cuánto cuestan los zapatos ahora?

Toplak, West y Stanovich (2014):

- Si John puede beber un barril de agua en 6 días y Mary puede beber un barril de agua en 12 días, ¿cuánto tiempo necesitarán para beber un barril de agua juntos?
Nota. Desarrollado por Frederick.
- Jerry tuvo la 15ª puntuación más alta de la clase y la 15ª puntuación más baja. ¿Cuántos estudiantes hay en la clase?
Nota. Desarrollado por Frederick
- Un hombre compra un cerdo por 60\$, lo vende por 70\$, lo vuelve a comprar por 80\$ y finalmente lo vende por 90\$. ¿Cuánto ha ganado?
Nota. Adaptado de Dominowski (1994).
- Simon decidió invertir 8,000 dólares en el mercado de valores un día antes de 2008. Seis meses después de la inversión, el 17 de julio, las acciones que había comprado bajaron un 50%. Afortunadamente para Simon, del 17 de julio al 17 de octubre, las acciones que había comprado subieron un 75%. En este punto, Simon: a) está en quiebra incluso en el mercado de valores, b) ha obtenido beneficios, c) ha perdido dinero.

Baron, Scott, Fincher y Metz (2015):

- En un lateral de un barco cuelga una escalera con seis peldaños. Cada peldaño está a un pie del siguiente y el peldaño inferior descansa sobre la superficie del agua. La marea sube a una velocidad de un pie por hora. ¿Cuánto tardará el agua en llegar al peldaño superior? a) 5 horas, b) 6 horas, c) nunca.

Nota. Cedido por Edward Royzman en una comunicación personal con los autores.

Reuben, Sapienza y Zingales (2015):

- Si lanzas una moneda normal tres veces, ¿cuál es la probabilidad de que salga cara al menos una vez?

Nota. Ítem sugerido por Frederick.

- Dos coches están en curso de colisión, viajando uno hacia el otro en el mismo carril. El coche A está viajando a 70 millas por hora. El coche B está viajando a 80 millas por hora. ¿A qué distancia se encuentran ambos coches 1 min antes de chocar?

Nota. Ítem sugerido por Frederick.

Primi, Morsanyi, Chiesi, Donati y Hamilton (2015):

- Si tres elfos pueden envolver tres juguetes por hora, ¿cuántos elfos se necesitarán para envolver seis juguetes en 2 horas?
- En un equipo de atletismo, los deportistas altos tienen tres veces más probabilidad de ganar una medalla que los deportistas bajos. Este año el equipo ha ganado 60 medallas hasta el momento. ¿Cuántas medallas han sido ganadas por los atletas bajos?

Nota. Los autores proponen en el apéndice del artículo 9 ítems adicionales considerados durante el desarrollo de la nueva escala.

Langnickel y Zeisberger (2016):

- Dos placas de hierro pesan 5.5 kg en total. La placa más pequeña pesa 5 kg menos que la más grande. ¿Cuántos kg pesa la placa pequeña?
- 3 trabajadores necesitan 3 días para producir 3 guitarras. ¿Cuántos días necesitarían 10 trabajadores para producir 10 guitarras?
- En un prado, hay una zona llena de flores. Cada día, la zona de flores duplica su tamaño. Si las flores tardan 48 días en cubrir todo el prado, ¿cuántos días tardarían las flores en cubrir la mitad de la pradera?

Thomson y Oppenheimer (2016):

- Si está corriendo en una carrera y rebasa a la persona que va en segundo lugar, ¿en qué lugar se encuentra usted?
- Un granjero tenía 15 ovejas y todas menos 8 murieron. ¿Cuántas ovejas le quedan?
- El padre de Emily tiene 3 hijas. Las dos primeras se llaman Abril y Mayo. ¿Cómo se llama la tercera hija?
- ¿Cuántos metros cúbicos de tierra hay en un agujero de 3 metros de profundo x 3 metros de ancho x 3 metros de largo?



APÉNDICE C.

Distribución de las Puntuaciones en las Nuevas Versiones del TRC

Autores	N.º ítems	Media	SD	N	r_{xx}	Observaciones
Finucane y Gullion (2010)	6 ítems (3 originales + 3 nuevos)	2.50	2.10	205	.82	Edad 25-45
		2.10	2.00	208	.80	Edad 65-74
Kinnunen y Windmann (2013)	4 ítems (3 originales + 1 nuevo)	1.70	1.80	195	.77	Edad 75-97
Borghans y Golsteyn (2014)	8 ítems (3 originales + 5 nuevos)	-	-	48	.69	
Del Missier, Visentini y Mantyla (2014)	6 ítems (Finucane y Gullion, 2010)	3.31	2.24	4,312	-	
Grossman, van der Weele y Andrijevnik (2014)	4 ítems (1 de Mata et al., 2013 + 2 originales + 1 nuevo)	1.60	-	224	-	Manipulación carga cognitiva
Morsanyi, Busdraghi y Primi (2014)	8 ítems (3 originales + 5 nuevos)	-	-	89	.69	Experimento 3
Shulman y McCallum (2014)	5 ítems (3 originales + 2 nuevos)	2.6	1.50	184	-	
Toplak, West y Stanovich (2014)	7 ítems (3 originales + 4 nuevos)	1.47	1.71	160	.72	Total 7 ítems
Baron, Scott, Fincher y Metz (2015)	6 ítems (Finucane y Gullion, 2010)	.98	1.07	160	-	Solo 4 ítems nuevos
	5 ítems (3 originales + 2 nuevos)	3.79	-	103	-	Estudio 1
	Finucane y Gullion, 2010)	2.78	-	82	-	Estudio 2
	7 ítems (Toplak et al., 2014)	2.66	-	104	.89	Estudio 3
		2.78	-	101	-	Estudio 5
Calvillo y Burgeno (2015)		0.98	1.30	184	.60	Estudio 1
		-	-	306	.71	Estudio 2

Continúa

Continuación

Autores	N.º ítems	Media	SD	N	r_{xx}	Observaciones
Corgnet, Espín, Hernán-González (2015)	7 ítems (Toplak et al., 2014)	3.67	2.25	83	-	Estudio 1. Hombres
		2.39	1.95	67	-	Estudio 1. Mujeres
		3.22	1.73	76	-	Estudio 2. Hombres
		2.18	1.35	82	-	Estudio 2. Mujeres
Corgnet, Espín, Hernán-González, Kujal y Rassenti (2015)	7 ítems (Toplak et al., 2014)	4.26	-	53	.72	Estudio 2. Hombres
		3.06	-	47	-	Estudio 2. Mujeres
Pennycook, Cheyne, Barr, Kochler y Fugelsang (2015)	7 ítems (Toplak et al., 2014)	-	-	279	.74	Estudio 1
Primi, Morsanyi, Chiesi, Donati y Hamilton (2015)	6 ítems (3 originales + 3 nuevos)	3.04	-	988	.76	Estudio 2
		1.34	-	70	.79	Estudio 3. Adolescentes
		3.33	-	287	.68	Estudio 3. Adultos
Reuben, Sapienza y Zingales (2015)	4 ítems (2 originales + 2 nuevos)	2.47	1.29	284	-	
Royzman, Landy y Leeman (2015)	6 ítems (Finucane y Gullion, 2010)	3.60	2.19	548	.85	Estudio 1
		4.27	2.02	74	.85	Estudio 2
Corgnet, Espín y Hernán-González (2016)	7 ítems (Toplak et al., 2014)	3.53	2.26	150	-	
Jackson, Kleitman, Howie y Stankov (2016)	7 ítems (Toplak et al., 2014)	2.07	1.85	250	.70	
Ring, Neysc, David-Barett y Schmidt (2016)	7 ítems (Toplak et al., 2014)	4.44	1.84	131	-	10 min. para responder
Thomson y Oppenheimer (2016)	4 ítems (verbales)	-	-	134	.51	
	7 ítems (3 originales + 3 verbales)	-	-	134	.71	

Continúa

Continuación

Autores	N.º ítems	Media	SD	N	r_{xx}	Observaciones
Morsanyi, McCormack y O'Mahony (2017)	6 ítems (Primi et al., 2015)	2.03	1.44	69	.67	
Primi, Morsanyi, Donati, Galli y Chiesi (2017)	6 ítems (Primi et al., 2015)	3.79	1.74	560	.76	
Toplak, West y Stanovich (2017)	11 ítems (3 originales + 4 Toplak et al., 2014 + 4 nuevos)	5.77	3.02	232	.80	
Yılmaz y Saribay (2017)	4 ítems (verbales; Thomson y Oppenheimer, 2016)	-	-	426	.64	
Primi, Donati, Chiesi y Morsanyi (2018)	6 ítems (Primi et al., 2015)	-	-	838	.80	Estudio 1
		2.82	1.86	181	.76	Estudio 2
Estudio 3 de esta tesis	13 ítems (3 originales + 10 Salgado, 2014a)	7.20	3.29	1,367	.78 ^a	

$$\bar{r}_{xx} = .74$$

Nota. SD = desviación típica; N = tamaño de la muestra de cada estudio; r_{xx} = coeficiente de fiabilidad por consistencia interna Alfa de Cronbach; \bar{r}_{xx} = coeficiente de fiabilidad alfa promedio.

^a Coeficiente de fiabilidad de equivalencia y estabilidad (CES). No integrado en la fiabilidad promedio.



APÉNDICE D.

TRC-13 (Salgado, 2014a) y Estadísticos Descriptivos del Test

Tabla D1.

TRC-13 Definitivo (Salgado, 2014a)

N.º ítem	Denominación ítem abreviada	Ítem	Opciones respuesta
1	Bate y pelota	Un bate y una pelota cuestan 1.10 euros. El bate cuesta un euro más que la pelota. ¿Cuánto cuesta la pelota?	R. intuitiva: 10 cént. R. correcta: 5 cént.
2	Máquinas	Si 5 máquinas tardan 5 minutos en hacer 5 aparatos, ¿cuánto tiempo tardan 100 máquinas en hacer 100 aparatos?	R. intuitiva: 100 min R. correcta: 5 min.
3	Nenúfares	En un lago, hay una zona de hojas de nenúfar. Cada día, la zona se duplica en tamaño. Si se tardan 48 días para que la zona cubra todo el lago, ¿cuánto tiempo haría falta para que la zona de nenúfares cubra la mitad del lago?	R. intuitiva: 24 días R. correcta: 47 días
4	Veleros	18 veleros navegando a 18 nudos tardan 18 días en ir del puerto A al puerto B. ¿Cuánto tardarán 36 veleros a 36 nudos?	R. intuitiva: 36 días R. correcta: 9 días
5	Veinteava parte	La veinteava (1/20) parte de veinte piezas de una cosa es igual a 2. ¿A cuánto será igual la cuarentava (1/40) parte de cuarenta piezas de la misma cosa?	R. intuitiva: 4 R. correcta: 2
6	Plátano y pastel	Un plátano y un pastel cuestan 37 céntimos. El plátano cuesta 13 céntimos más que el pastel. ¿Cuánto cuesta el pastel?	R. intuitiva: 24 cént. R. correcta: 12 cént.

Continúa

Tabla D1.

Continuación

N.º ítem	Denominación ítem abreviada	Ítem	Opciones respuesta
7	Taxista	La probabilidad de que un taxista en una gran capital tenga un accidente al mes es del 10%. ¿Cuál será la probabilidad de que tenga 1 accidente en dos meses seguidos, por tanto, dos accidentes en total?	R. intuitiva: 20% R. correcta: 1%
8	Posibilidad de aprobar	En un colegio, los estudiantes asignados a la clase A aprueban todos, los asignados a la clase B aprueban la mitad y los asignados a la clase C aprueban uno de cada tres. ¿Qué posibilidades de aprobar tiene un nuevo estudiante antes de ser asignado a la clase A, B o C?	R. intuitiva: No se sabe antes de asignarlo R. correcta: 60%
9	Aviones	1,000 aviones tardan 1,000 minutos en transportar 1,000 pasajeros de la ciudad A a la ciudad B que está a más de 10,000 kilómetros de distancia. ¿Cuántos minutos tardarán 10,000 aviones en transportar 10,000 pasajeros?	R. intuitiva: 10,000 min. R. correcta: 1,000 min.
10	Máquinas coste pieza	Si una máquina produce una pieza por 100 céntimos y 50 máquinas producen 50 piezas por 50 céntimos, ¿cuántas máquinas se necesitarán para que el precio de las piezas sea 12.5 céntimos?	R. intuitiva: 125 R. correcta: 200
11	Hombres altos	En una ciudad, de cada 4 personas 3 son hombres y 2 de cada 4 hombres son altos. ¿Qué posibilidades hay de que un día al salir a la calle, la primera persona que veamos sea un hombre alto?	R. intuitiva: 75% R. correcta: 37.5%
12	Sillas	¿De cuántos modos diferentes pueden sentarse 6 personas en una mesa con 6 sillas?	R. intuitiva: 36 R. correcta: 720
13	X + Y	$X + Y = 110$. $X = 100 + Y$. ¿Cuánto vale Y?	R. intuitiva: 10 R. correcta: 5

Tabla D2.
Estadísticos Descriptivos del TRC-13 (Definitivo)

	N Total	Hombres	Mujeres
Tamaño de la muestra	1,367	490	877
N.º ítems	13	13	13
Rango posibles puntuaciones	0-13	0-13	0-13
Rango puntuaciones alcanzadas	0-13	0-13	0-13
Puntuación Promedio			
Media	7.20	8.46	6.49
Mediana	7	9	6
Moda	10	11	5
Medidas de variación			
Desviación Estándar	3.29	3.08	3.19
Asimetría	-.18	-.63	.03
Curtosis	-0.86	-0.30	-0.82
Cuartiles			
25	5	6	4
50	7	9	6
75	10	11	9

Nota. N total = muestra total.



APÉNDICE E.

Medias, Desviaciones Típicas y Diferencias en el Desempeño en el TRC según la Condición Experimental (Hombres)

Desempeño en el test					Diferencias en el desempeño				
Condición 1		Condición 2		Condición 3	Condición 1 vs. 2		Condición 1 vs. 3		Condición 2 vs. 3
\bar{X}	SD	\bar{X}	SD	\bar{X}	d	F	d	F	F
Ítem 1	.45	.50	.48	.52	.50	.43	6.47*	-.14	.83
Ítem 2	.57	.50	.44	.72	.45	.38	5.31*	-.32	4.15*
Ítem 3	.55	.50	.45	.55	.50	.36	4.34*	.00	.00
Ítem 4	.83	.38	.43	.81	.39	.17	.79	.05	.04
Ítem 5	.68	.47	.46	.67	.47	.07	.10	.02	.01
Ítem 6	.78	.42	.36	.79	.41	.18	1.13	-.02	.02
Ítem 7	.67	.48	.44	.62	.49	.15	.76	.10	.42
Ítem 8	.55	.50	.50	.50	.50	.08	.18	.10	.39
Ítem 9	.83	.38	.34	.83	.38	.11	.45	.00	.00
Ítem 10	.54	.50	.49	.55	.50	.12	.62	-.02	.02
Ítem 11	.88	.32	.21	.83	.38	.30	2.40	.14	1.03
Ítem 12	.57	.50	.50	.50	.50	.00	.00	.14	.65
Ítem 13	.59	.50	.50	.50	.50	.12	.58	.18	1.36
TRC-13	8.48	3.34	2.73	8.38	2.90	-.25	2.11	.03	.04
TRC-3	1.57	1.11	0.96	1.79	1.02	-.54	10.27**	-.21	1.74
TRC-10	6.91	2.48	2.02	6.59	2.21	-.08	.24	.14	.72

Nota: Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; SD = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = sumatorio de los ítems originales (Frederick, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el ítem 10.

† $p \leq .10$; * $p \leq .05$; ** $p \leq .01$



APÉNDICE F.

Medias, Desviaciones Típicas y Diferencias en el Desempeño en el TRC según la Condición Experimental (Mujeres)

Desempeño en el test						Diferencias en el desempeño											
Condición 1			Condición 2			Condición 3			Condición 1 vs. 2			Condición 1 vs. 3			Condición 2 vs. 3		
\bar{X}	SD		\bar{X}	SD		\bar{X}	SD		d	F	d	F	d	F	d	F	
Ítem 1	.25	.44	.45	.50		.42	.50		-.42	11.84***	-.36	7.96**	.06		.29		
Ítem 2	.48	.50	.67	.47		.57	.50		-.39	10.42***	-.18	2.02	.21		2.95 [†]		
Ítem 3	.39	.49	.46	.50		.31	.46		-.14	1.20	.17	1.84	.31		6.01*		
Ítem 4	.70	.46	.70	.46		.59	.49		.00	.00	.23	3.40 [†]	.23		3.39 [†]		
Ítem 5	.50	.50	.56	.50		.54	.50		-.12	1.09	-.08	.43	.04		.14		
Ítem 6	.59	.49	.63	.49		.63	.48		-.08	.43	-.08	.41	.00		.00		
Ítem 7	.46	.50	.56	.50		.38	.49		-.20	2.41	.16	1.53	.36		7.77**		
Ítem 8	.45	.50	.40	.49		.37	.48		.10	.62	.16	1.58	.06		.24		
Ítem 9	.80	.40	.84	.37		.72	.45		-.10	.80	.19	2.17	.29		5.62*		
Ítem 10	.42	.50	.46	.50		.38	.49		-.08	.34	.08	.38	.16		1.43		
Ítem 11	.63	.48	.72	.45		.64	.48		-.19	2.12	-.02	.02	.17		1.66		
Ítem 12	.41	.49	.37	.49		.42	.50		.08	.43	-.02	.02	-.10		.47		
Ítem 13	.45	.50	.39	.49		.37	.48		.12	1.08	.16	1.91	.04		.14		
TRC-13	6.54	3.22	7.21	2.93		6.33	2.98		-.22	3.04 [†]	.07	.27	.30		5.48*		
TRC-3	1.12	1.10	1.58	1.01		1.29	.96		-.44	12.51***	-.16	1.77	.29		5.39*		
TRC-10	5.42	2.44	5.63	2.20		5.04	2.34		-.09	.50	.16	1.57	.26		4.17*		

Nota. Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; *SD* = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10; TRC-3 = sumatorio de los ítems originales (Frederiek, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el ítem 10.

† $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$



APÉNDICE G.

Medias y Desviaciones Típicas de los Tiempos de Respuesta (en Segundos) del TRC según la Condición Experimental (Hombres)

	Tiempo de Respuesta (TR)				Diferencias en el TR			
	Condición 1		Condición 2		Condición 1 vs. 2		Condición 1 vs. 3	
	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD
Ítem 1	31.97	25.28	27.69	17.69	33.08	29.85	.20	1.32
Ítem 2	45.03	38.40	32.78	29.61	30.84	26.09	.36	4.36*
Ítem 3	45.33	30.11	40.23	19.64	41.63	29.62	.20	1.37
Ítem 4	32.85	32.38	51.20	63.49	44.77	32.70	-.37	4.56*
Ítem 5	54.17	34.97	49.62	30.07	61.52	44.60	.14	.67
Ítem 6	56.58	46.00	64.95	52.66	69.78	49.48	-.17	.98
Ítem 7	41.67	29.90	40.18	20.58	51.27	33.35	.06	.12
Ítem 8	67.45	63.39	68.82	48.70	72.53	46.57	-.02	.02
Ítem 9	36.38	21.01	39.64	22.39	46.39	30.27	-.15	.78
Ítem 10	72.95	54.71	62.12	40.70	74.26	64.24	.22	1.72
Ítem 11	57.85	46.76	55.45	31.16	64.29	43.13	.06	.12
Ítem 12	35.72	27.28	48.28	37.13	46.35	36.95	-.39	5.10*
Ítem 13	89.50	74.50	93.25	73.35	87.63	68.31	-.05	.09
TRC-13	667.44	284.88	674.23	249.15	724.35	321.43	-.03	.02
TRC-3	122.32	71.41	100.70	44.10	105.55	62.52	.36	4.53*
TRC-10	545.11	237.28	573.53	220.43	618.79	273.85	-.12	.53

Nota. Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; SD = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = sumatorio de los ítems originales (Frederick, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el ítem 10.

† $p \leq .10$; * $p \leq .05$; ** $p \leq .01$



APÉNDICE H.

Medias y Desviaciones Típicas de los Tiempos de Respuesta (en Segundos) del TRC según la Condición Experimental (Mujeres)

	Tiempo de Respuesta (TR)				Diferencias en el TR			
	Condición 1		Condición 2		Condición 3		Condición 1 vs. 2	
	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	d	F
Ítem 1	31.49	26.98	29.16	21.77	32.67	25.42	.10	.58
Ítem 2	36.63	24.44	31.67	25.81	29.17	26.00	.20	2.52
Ítem 3	37.64	22.99	40.83	20.39	39.19	26.76	-.15	1.40
Ítem 4	31.24	28.13	46.87	45.87	46.87	39.75	-.41	10.87**
Ítem 5	49.26	41.82	50.16	36.95	52.89	47.15	-.02	.03
Ítem 6	49.81	38.39	66.55	67.59	60.80	45.04	-.30	5.97*
Ítem 7	41.73	32.58	39.72	23.80	42.92	29.84	.07	.32
Ítem 8	57.54	46.66	55.04	43.82	58.52	43.43	.06	.20
Ítem 9	34.83	22.18	37.01	19.66	40.23	25.87	-.10	.70
Ítem 10	65.29	48.20	66.42	47.61	70.19	49.79	-.02	.04
Ítem 11	69.70	51.43	68.00	49.27	69.28	46.62	.03	.07
Ítem 12	34.38	29.16	33.78	25.81	41.44	32.26	.02	.03
Ítem 13	94.91	79.30	90.52	82.07	87.72	65.77	.05	.19
TRC-13	634.46	306.70	655.74	284.78	671.90	272.96	-.07	.34
TRC-3	105.76	51.02	101.66	46.67	101.03	55.29	.08	.45
TRC-10	528.71	275.10	554.08	256.09	570.86	238.83	-.10	.59

Nota. Condición 1 = orden 1-13; Condición 2 = orden 4-9, 1-3 y 10-13; Condición 3 = 4-13 y 1-3; \bar{X} = media; SD = desviación típica; TRC-13 = total ítems TRC-3 y TRC-10 Salgado (2014a); TRC-3 = sumatorio de los ítems originales (Frederick, 2005); TRC-10 = sumatorio de los ítems desarrollados por Salgado (2014a) eliminando el ítem 10.

† $p \leq .10$; * $p \leq .05$; ** $p \leq .001$

APÉNDICE I.

Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables (Muestra Total)

	\bar{X}	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Edad	21.72	4.14	-													
2. TRC-13	7.20	3.29	-.07**	(.66)												
3. TRC-3	1.42	1.08	-.03	.82**	(.71)											
4. TRC-10	5.78	2.48	-.07*	.97**	.65**	(.57)										
5. WPT	22.36	5.22	-.04	.48**	.41**	.46**	(.85)									
6. RLS	13.67	3.76	-.18**	.35**	.35**	.33**	.52**	(.69)								
7. FACTOR_G	23.18	4.63	-.12*	.39**	.34**	.38**	.30**	.34**	(.71)							
8. Autocontrol	37.67	8.77	.15**	-.04	-.04	-.03	-.04	-	-.14	(.84)						
9. Impulsividad	36.99	5.68	-.09**	.02	.05	.00	.10**	.07	.20*	-.29**	(.63)					
10. Atención	35.71	6.47	.05	.13**	.10**	.12**	.12**	.23**	-.08	.40**	-.07	(.71)				
11. Esfuerzo	36.64	7.57	.07*	.03	.01	.04	.00	.19*	-.06	.65**	-.13**	.58**	(.81)			
12. Intuición	41.03	5.88	.00	.03	.02	.04	.08*	.08	.06	-.07	.49**	.14**	.05	(.77)		
13. Riesgo	35.98	5.73	-.02	.12**	.09*	.12**	.07	-.05	.07	-.29**	.41**	-.02	-.13**	.33**	(.66)	
14. NFC	36.58	5.50	.02	.25**	.23**	.23**	.26**	.21**	.18*	.18**	.15**	.39**	.33**	.36**	.21**	(.75)
15. TFE	12.17	4.01	-.16*	.34**	.30**	.31**	.32**	-	-	.01	-.02	.11	.01	.09	.09	.13
16. EE_IP	35.00	14.98	.16**	.14**	.13**	.13**	.20**	.12*	-.12*	.23**	.07	.24**	.22**	.18**	.24**	.27**
17. EX_IP	52.34	11.05	.05	.04	.06	.03	.09*	-.01	.04	.14*	.41**	.04	.10*	.42**	.36**	.13*
18. AP_IP	53.01	10.63	.04	.11**	.12**	.10*	.12**	.06	.05	.11	.26**	.21**	.19**	.41**	.39**	.50**
19. AM_IP	40.52	9.48	.16**	-.03	.01	-.04	-.04	-.07	-.08	.14*	-.13*	-.07	-.08	-.08	-.03	-.08
20. C_IP	35.98	11.67	.06	-.06	-.05	-.05	.03	.10	-.11	.54**	-.15**	.38**	.58**	-.05	-.42**	.14**
21. EE_Q15F	22.67	6.62	.06	.05	-.01	.06	.04	.02	.06	.04	-.07*	.10**	.01	.00	.11**	.04
22. EX_Q15F	28.00	8.04	-.03	-.03	-.05	-.02	-.03	-.09	.04	-.08	.40**	-.04	-.02	.27**	.32**	.05
23. AP_Q15F	28.83	9.54	.03	.16**	.16**	.14**	.15**	.05	.17*	-.20**	.08*	.09*	-.01	.15**	.21**	.50**
24. AM_Q15F	29.59	6.81	.09*	-.02	.05	-.05	-.10**	-.05	-.03	-.11*	-.26**	-.08*	-.16**	-.15**	-.01	-.10**
25. C_Q15F	26.22	8.22	.04	-.09**	-.09*	-.08*	-.10**	.10	-.13	.42**	-.36**	.16**	.34**	-.32**	-.54**	-.16**
26. GCT	26.95	11.28	-.03	.30**	.19**	.32**	.46**	-	-	-.01	.10*	.11*	.03	.07	.02	.12**
27. CDTE	103.11	15.09	.09**	-.10**	-.07*	-.11**	-.05	.09	-.08	.45**	-.14**	.27**	.50**	-.04	-.26**	.12**
28. CDCE	110.47	13.00	.02	-.07*	-.04	-.08*	-.06	.03	-.10	.24**	-.07	.20**	.34**	.09**	-.12**	.15**
29. CDAN	61.45	13.38	-.06	-.03	-.05	-.02	-.02	-.24**	-.14	-.41**	.25**	-.17**	-.38**	.13**	.27**	-.06
30. GPA	7.01	.88	-.03	.15**	.16**	.14**	.17**	.21**	.14	.24**	-.11**	.13**	.21**	-.08*	-.15**	.16**
31. Bachiller	7.66	1.22	-.13	.35**	.31**	.33**	.41**	-	-	.09	.04	.03	.08	-.13	.03	.11
32. CAU	7.19	1.20	-.20**	.36**	.31**	.35**	.41**	-	-	.02	.06	.02	.05	-.15*	.01	.13

Nota. $N = 1,367 - 125$. TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); TRC-10 = ítems Salgado (2014a); TRC-3 = ítems figuras enmascaradas; EE_IP = estabilidad emocional medida con el inventario IP/5F; EX_IP = extraversión medida con el inventario IP/5F; medido con el inventario IP/5F; EE_Q15F = estabilidad emocional medida con el inventario quasi-ipsativo Q15F; EX_Q15F = extroversión medida con el inventario quasi-ipsativo Q15F; C_Q15F = conciencia medida con el inventario quasi-ipsativo Q15F; GCT = general contraproductivas; GPA = grade point average; CAU = calificaciones de acceso a la universidad.

* $p < .05$; ** $p < .01$

15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
(.73)																	
-	(.91)																
-	.36**	(.85)															
-	.21**	.46**	(.84)														
-	.12**	.04	.02	(.78)													
-	.01	-.08*	.01	-.11**	(.85)												
.00	.58**	-.02	-.10	-.01	-.12*	(.66)											
-.06	.18**	.67**	.22**	-.08	-.22**	-.03	(.75)										
.13*	-.04	-.07	.51**	-.14**	-.01	.12**	-.15**	(.83)									
.00	-.03	-.17**	-.12*	.63**	-.22**	-.02	-.19**	-.15**	(.69)								
-.05	-.29**	-.40**	-.31**	-.10	.61**	-.19**	-.44**	-.15**	-.14**	(.77)							
.19**	-	-	-	-	-	-.01	-.04	.05	-.12*	-.02	(.78)						
.01	-.01	.02	.07	.07	.53**	-.15**	-.10**	-.03	-.05	.39**	.07	(.89)					
-.07	.07	.23**	.20**	.14**	.39**	-.12**	.03	-.03	.07*	.17**	-.03	.61**	(.86)				
-.08	-.05	.07	-.04	-.15**	-.35**	.09*	.15**	.07*	-.04	-.31**	-.03	-.56**	-.39**	(.90)			
.12	.07	-.08	.00	-.06	.18**	-.09*	-.07	.07	-.09**	.22**	.06	.26**	.17**	-.24**	(.87)		
-	-	-	-	-	-	-.04	.03	.10	-.16	-.04	.21*	.04	-.09	-.02	.43**	-	
-	-	-	-	-	-	-.05	.01	.10	-.07	-.04	.27**	.10	-.04	-.07	.40**	.97**	-

Frederick (2005); WPT = Wonderlic personnel test; RLS = test de razonamiento lógico simbólico; NFC = need for cognition; TFE = test de AP_IP = apertura a la experiencia medida con el inventario IP/5F; AM_IP = amigabilidad medido con el inventario IP/5F; C_IP = conciencia con el inventario quasi-ipsativo QISF; AP_QISF= apertura a la experiencia medido con el inventario quasi-ipsativo QISF; AM_QISF = clerical test; CDTE = escala de desempeño de tarea; CDCE = escala de desempeño contextual; CDAN = escala de conductas

APÉNDICE J.

Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables (Hombres)

	\bar{X}	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Edad	22.06	4.85	-													
2. TRC-13	8.46	3.08	-.13**	.67												
3. TRC-3	1.72	1.08	-.03*	.81**	.76											
4. TRC-10	6.74	2.29	-.13**	.96**	.62**	.52										
5. WPT	23.77	5.55	-.02	.43**	.40**	.41**	.84									
6. RLS	14.58	3.62	-.15	.24**	.33**	.17	.53**	.74								
7. FACTOR_G	22.46	5.07	-.08	.49**	.44**	.45**	.11	.07	.76							
8. Autocontrol	38.51	8.64	.13*	-.03	-.06	-.01	.04	-	-.11	.84						
9. Impulsividad	37.21	5.53	-.11	.05	.04	.05	.16*	-.01	.39**	-.20*	.62					
10. Atención	36.22	6.42	.03	.12	.15*	.10	.05	.12	-.25	.47**	-.13*	.70				
11. Esfuerzo	36.90	7.22	.07	.08	.04	.09	.04	.24*	-.16	.63**	-.17**	.59**	.79			
12. Intuición	41.31	5.35	-.04	.06	.06	.05	.09	-.04	.07	-.06	.40**	.15*	.06	.73		
13. Riesgo	36.66	5.84	-.02	.19**	.14*	.18**	.01	-.22	.36*	-.17*	.39**	.00	-.11	.31**	.67	
14. NFC	37.81	5.22	-.08	.26**	.23**	.25**	.24**	.11	.13	.31**	.12	.42**	.39**	.32**	.12*	.73
15. TFE	11.86	4.30	-.19	.35**	.38**	.29*	.31*	-	-	.14	.24	.14	.01	.15	.16	.11
16. EE_IP	42.37	13.80	.21**	.07	.06	.07	.19*	.12	-.04	.33**	.01	.24**	.42**	.12	.07	.18*
17. EX_IP	53.44	10.68	.08	.04	.09	.01	.11	.02	.10	.28**	.39**	.04	.18*	.47**	.29**	.08
18. AP_IP	53.87	10.76	.06	.11	.12	.10	.10	-.02	.12	.19*	.23**	.23**	.31**	.45**	.39**	.48**
19. AM_IP	40.33	10.12	.16**	-.12	.02	-.17**	-.04	-.08	-.17	.09	-.08	-.06	-.06	-.01	-.01	-.07
20. C_IP	34.77	11.60	.08	-.09	-.13	-.06	.06	.17	-.18	.55**	-.24**	.45**	.60**	-.02	-.47**	.15
21. EE_Q15F	25.01	6.72	.06	.07	-.02	.11	.03	.04	.21	.01	-.11	.06	.02	-.06	.11	.02
22. EX_Q15F	28.40	7.97	-.11	.06	.01	.08	.05	-.16	.21	-.07	.40**	.02	.02	.30**	.30**	.07
23. AP_Q15F	29.79	9.86	-.05	.13*	.12*	.12	.13*	-.12	.12	-.10	.11	.16*	.07	.21**	.26**	.51**
24. AM_Q15F	29.21	7.32	.16**	-.04	.00	-.05	-.06	.05	-.10	-.19*	-.23**	-.06	-.13*	-.15*	-.01	-.09
25. C_Q15F	24.31	7.99	.08	-.20**	-.17**	-.19**	-.16*	.08	-.33**	.44**	-.39**	.15*	.35**	-.31**	-.56**	-.13
26. GCT	27.13	12.70	-.10	.34**	.30**	.34**	.47**	-	-	-.04	.04	.07	.03	-.04	-.10	.22*
27. CDTE	100.31	15.95	.11*	-.11	-.10	-.10	.01	.19	-.22	.45**	-.22**	.26**	.51**	-.14*	-.31**	.07
28. CDCE	108.72	13.29	.03	-.02	-.02	-.02	.01	.09	-.27	.27**	-.15*	.22**	.36**	.01	-.16*	.10
29. CDAN	63.43	14.50	-.06	-.07	-.07	-.07	-.08	-.25*	.07	-.36**	.23**	-.16*	-.39**	.11	.17**	-.05
30. GPA	6.97	.92	.01	.08	.12	.06	.15*	.19	-.11	.25**	-.12	.08	.26**	-.16*	-.15*	.08
31. Bachiller	7.71	1.29	-.21	.36*	.47**	.29	.34*	-	-	.33*	-.11	.01	.16	-.28	-.16	.09
32. CAU	7.21	1.20	-.22	.42**	.45**	.38**	.35**	-	-	-.03	.08	-.07	-.01	-.14	.06	-.01

Nota. $N = 1,367 - 125$. TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); TRC-10 = ítems Salgado (2014a); TRC-3 = ítems figuras enmascaradas; EE_IP = estabilidad emocional medida con el inventario IP/5F; EX_IP = extraversión medida con el inventario IP/5F; medido con el inventario IP/5F; EE_Q15F = estabilidad emocional medida con el inventario quasi-ipsativo Q15F; EX_Q15F = extraversión medida con el inventario quasi-ipsativo Q15F; C_Q15F = conciencia medida con el inventario quasi-ipsativo Q15F; GCT = general contraproductivas; GPA = grade point average; CAU = calificaciones de acceso a la universidad;

* $p < .05$; ** $p < .01$

15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
.73																	
-	.89																
-	.39**	.84															
-	.20**	.47**	.85														
-	.10	.07	.01	.81													
-	.12	-.06	.09	-.12	.85												
.22	.51**	-.22*	-.25**	-.11	-.04	.65											
.17	.28**	.72**	.28**	-.08	-.21*	-.08	.75										
-.04	-.04	-.12	.53**	-.14	.01	-.08	-.17**	.84									
.02	-.10	-.16	-.12	.72**	-.18*	-.09	-.19**	-.16**	.81								
-.27*	-.09	-.32**	-.28**	-.15	.66**	-.10	-.40**	-.18**	-.17**	.85							
.09	-	-	-	-	-	.07	-.08	.11	.02	-.04	.96						
-.05	.15	-.04	.14	.08	.59**	-.15*	-.11	.01	.02	.42**	.09	.90					
.03	.16*	.15*	.18*	.06	.45**	-.14*	.08	.03	.14*	.18**	-.02	.61**	.87				
.06	-.20*	.00	-.04	-.15	-.32**	.04	.03	.06	-.03	-.20**	-.05	-.53**	-.43**	.91			
.06	.13	-.01	-.01	-.07	.16*	-.14*	-.06	.06	-.09	.24**	.05	.35**	.26**	-.32**	.87		
-	-	-	-	-	-	-.09	-.03	.05	-.15	.05	.36*	.18	-.17	-.23	.46**	-	
-	-	-	-	-	-	-.19	.08	.12	-.05	-.11	.37**	.13	-.16	-.19	.44**	.95**	-

Frederick (2005); WPT = Wonderlic personnel test; RLS = test de razonamiento lógico simbólico; NFC = need for cognition; TFE = test de AP_IP = apertura a la experiencia medida con el inventario IP/5F; AM_IP = amigabilidad medido con el inventario IP/5F; C_IP = conciencia con el inventario quasi-ipsativo QISF; AP_QISF = apertura a la experiencia medido con el inventario quasi-ipsativo QISF; AM_QISF = clerical test; CDTE = escala de desempeño de tarea; CDCE = escala de desempeño contextual; CDAN = escala de conductas.

APÉNDICE K.

Medias, Desviaciones Típicas y Correlaciones Observadas entre las Variables (Mujeres)

	\bar{X}	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Edad	21.52	3.66	-													
2. TRC-13	6.49	3.19	-.07*	.61												
3. TRC-3	1.24	1.05	-.04	.81**	.69											
4. TRC-10	5.24	2.42	-.07*	.97**	.63**	.53										
5. WPT	21.65	4.89	-.08*	.46**	.38**	.45**	.84									
6. RLS	13.10	3.74	-.26**	.35**	.31**	.34**	.47**	.59								
7. FACTOR_G	23.71	4.21	-.17*	.45**	.33**	.45**	.42**	.46**	.65							
8. Autocontrol	37.19	8.81	.17**	-.08	-.06	-.09	-.10	-	-.15	.84						
9. Impulsividad	36.88	5.75	-.09*	-.01	.05	-.03	.07	.09	.09	-.33**	.64					
10. Atención	35.47	6.48	.05	.11**	.07	.12**	.14**	.28**	.01	.37**	-.04	.71				
11. Esfuerzo	36.51	7.73	.07	.00	-.01	.01	-.03	.16	-.02	.66**	-.12**	.57**	.81			
12. Intuición	40.90	6.11	.01	.01	-.01	.02	.07	.09	.07	-.08	.52**	.13**	.04	.79		
13. Riesgo	35.66	5.66	-.04	.06	.04	.07	.07	-.01	-.14	-.35**	.42**	-.03	-.15**	.34**	.65	
14. NFC	36.01	5.53	.05	.19**	.19**	.18**	.24**	.21*	.23*	.11*	.16**	.37**	.30**	.37**	.24**	.75
15. TFE	12.29	3.91	-.14	.37**	.29**	.36**	.34**	-	-	-.04	-.10	.11	.01	.08	.07	.14
16. EE_IP	30.28	13.77	.13**	-.01	-.01	-.02	.05	-.002	-.09	.12	.03	.24**	.16*	.14*	.25**	.18**
17. EX_IP	51.64	11.23	.03	.00	.01	.00	.04	-.09	.03	.01	.41**	.02	.07	.39**	.38**	.11
18. AP_IP	52.47	10.52	.03	.09	.11*	.07	.12*	.09	.01	.01	.26**	.20**	.13*	.38**	.37**	.51**
19. AM_IP	40.64	9.06	.16**	.04	.01	.05	-.03	-.07	.04	.21*	-.17**	-.07	-.10	-.12	-.05	-.09
20. C_IP	36.75	11.65	.05	.00	.03	-.01	.06	.11	-.08	.55**	-.07	.36**	.58**	-.03	-.37**	.19**
21. EE_Q15F	21.69	6.32	.02	-.06	-.08	-.05	-.01	-.01	-.01	.05	-.07	.11*	.00	.02	.10*	.01
22. EX_Q15F	27.83	8.06	.02	-.08*	-.08	-.07	-.08	-.08	-.03	-.09	.40**	-.06	-.03	.26**	.33**	.04
23. AP_Q15F	28.43	9.38	.06	.15**	.16**	.13**	.14**	.11	.20*	-.24**	.06	.06	-.04	.13**	.18**	.49**
24. AM_Q15F	29.75	6.58	.04	.00	.08	-.03	-.11**	-.11	.02	-.08	-.27**	-.08*	-.17**	-.16**	-.01	-.10*
25. C_Q15F	27.02	8.18	.05	.01	.00	.01	-.04	.14	-.04	.42**	-.34**	.17**	.35**	-.32**	-.53**	-.15**
26. GCT	26.88	10.68	.02	.29**	.13*	.32**	.46**	-	-	.01	.12*	.12*	.04	.11*	.07	.08
27. CDTE	104.46	14.48	.10**	-.05	-.01	-.06	-.05	.07	.00	.45**	-.09*	.29**	.51**	.01	-.22**	.17**
28. CDCE	111.32	12.79	.04	-.06	-.02	-.07	-.07	-.02	.00	.23**	-.02	.20**	.34**	.14**	-.10*	.21**
29. CDAN	60.52	12.73	-.07	-.07	-.09	-.06	-.02	-.27**	-.28**	-.45**	.25**	-.19**	-.38**	.13**	.31**	-.10*
30. GPA	7.03	.87	-.05	.21**	.20**	.19**	.20**	.23**	.25*	.24**	-.10*	.16**	.19**	-.05	-.14**	.21**
31. Bachiller	7.64	1.19	-.10	.34**	.24*	.36**	.44**	-	-	-.02	.11	.05	.05	-.07	.11	.12
32. CAU	7.18	1.21	-.19*	.34**	.27**	.34**	.44**	-	-	.04	.05	.05	.08	-.16	-.01	.18*

Nota. $N = 1,367 - 125$. TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a); TRC-10 = ítems Salgado (2014a); TRC-3 = ítems figuras enmascaradas; EE_IP = estabilidad emocional medida con el inventario IP/5F; EX_IP = extraversión medida con el inventario IP/5F; medido con el inventario IP/5F; EE_Q15F = estabilidad emocional medida con el inventario quasi-ipsativo Q15F; EX_Q15F = extroversión medida con el inventario quasi-ipsativo Q15F; C_Q15F = conciencia medida con el inventario quasi-ipsativo Q15F; GCT = general contraproductivas; GPA = grade point average; CAU = calificaciones de acceso a la universidad;

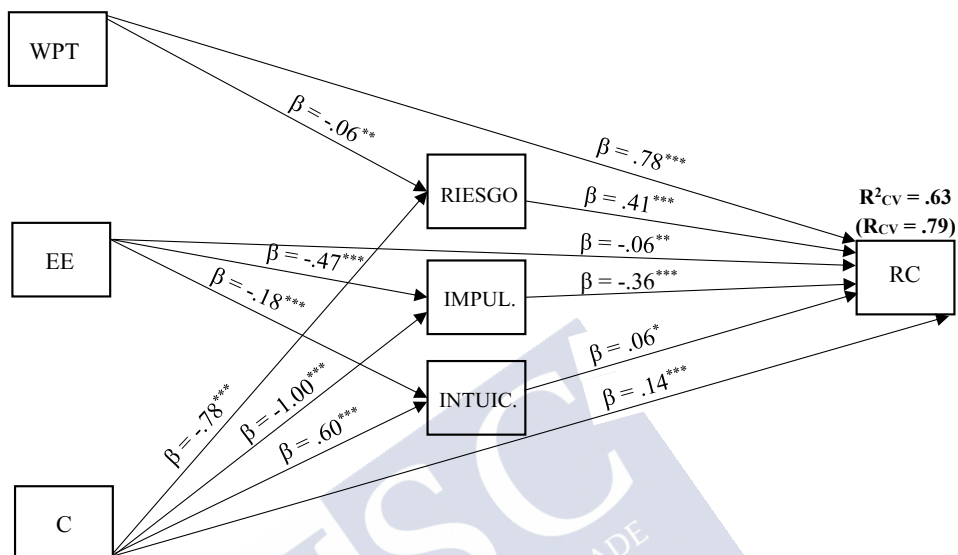
* $p < .05$; ** $p < .01$

15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
.73																	
-	.90																
-	.34**	.86															
-	.20**	.45**	.84														
-	.17**	.02	.03	.76													
-	.00	-.08	-.04	-.10*	.85												
-.09	.55**	.03	-.04	.04	-.11	.64											
-.14	.15*	.65**	.20**	-.08	-.23**	-.02	.76										
.21**	-.13*	-.07	.49**	-.13*	.01	-.16**	-.14**	.83									
-.01	-.02	-.19**	-.12	.58**	-.24**	.03	-.19**	-.14**	.68								
.03	-.29**	-.42**	-.33**	-.08	.57**	-.19**	-.45**	-.13**	-.13**	.77							
.24**	-	-	-	-	-	-.06	-.03	.02	-.18**	-.02	.71						
.03	.03	.09	.05	.06	.48**	-.10*	-.08*	-.04	-.10*	.35**	.07	.88					
-.11	.09	.30**	.22**	.19**	.35**	-.07	.02	-.04	.03	.14**	-.04	.60**	.86				
-.13	-.05	.11	-.05	-.14*	-.37**	.06	.20**	.06	-.03	-.34**	-.03	-.58**	-.36**	.89			
.15	.05	-.11*	.01	-.05	.19**	-.04	-.07	.07	-.10*	.20**	.07	.21**	.11**	-.20**	.87		
-	-	-	-	-	-	-.02	.05	.11	-.16	-.08	.13	-.02	-.06	.09	.42**	-	
-	-	-	-	-	-	.00	-.02	.09	-.08	-.02	.22**	.09	.00	-.02	.38**	.97**	-

Frederick (2005); WPT = Wonderlic personnel test; RLS = test de razonamiento lógico simbólico; NFC = need for cognition; TFE = test de AP_IP = apertura a la experiencia medida con el inventario IP/5F; AM_IP = amigabilidad medida con el inventario IP/5F; C_IP = conciencia con el inventario quasi-ipsativo QISF; AP_QISF = apertura a la experiencia medida con el inventario quasi-ipsativo QISF; AM_QISF = clerical test; CDTE = escala de desempeño de tarea; CDCE = escala de desempeño contextual; CDAN = escala de conductas.

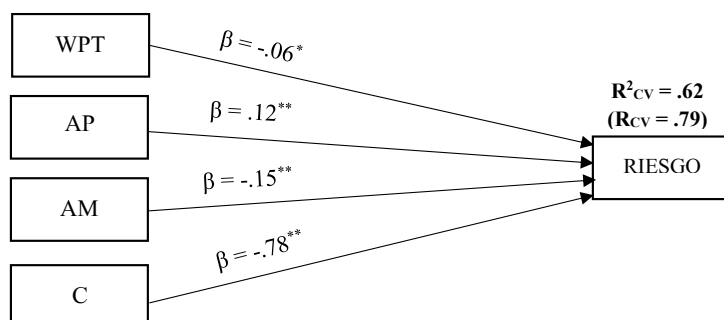
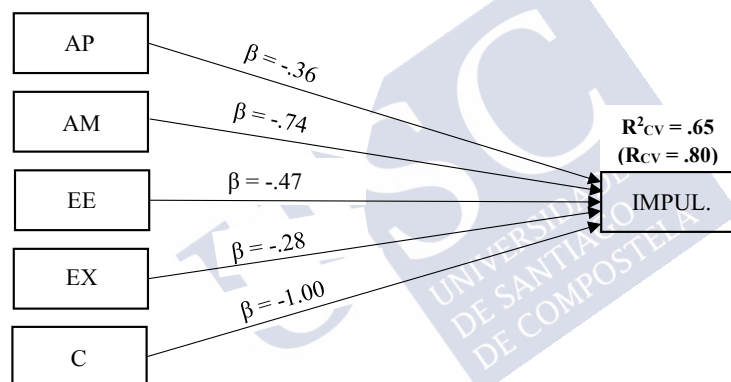
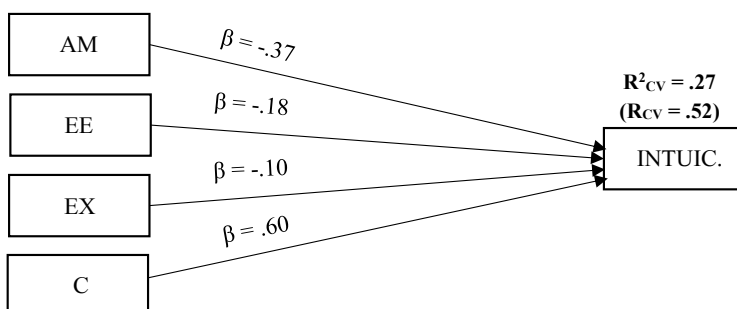


APÉNDICE L.

Efectos Directos en la Predicción de la RC (Muestra Total)Figura L. *Efectos Directos de las Variables en la Predicción de la RC (Muestra Total)*Nota. $*p < .05$; $**p < .01$; $***p < .001$



APÉNDICE M.

Efectos Indirectos en la Predicción de la RC (Muestra Total)Figura M1. *Efectos Indirectos en la Predicción de RC a través de Riesgo (Muestra Total)*Nota. * $p < .01$; ** $p < .001$ Figura M2. *Efectos Indirectos en la Predicción de RC a través de Impulsividad (Muestra Total)*Nota. Todos los pesos de regresión son significativos a nivel $p < .001$ Figura M3. *Efectos Indirectos en la Predicción de RC a través de Intuición (Muestra Total)*Nota. Todos los pesos de regresión son significativos a nivel $p < .001$



APÉNDICE N.

Predicción Múltiple del Desempeño en el TRC-13 en la Muestra de Hombres y Mujeres

Para efectuar los cálculos, se ha seguido el mismo procedimiento metodológico empleado en los análisis de la muestra total. Los análisis de regresión jerárquica se han realizado en tres pasos: (1) se ha introducido la variable de CMG (concretamente, el WPT) por ser esta la que presenta una mayor asociación teórica y empírica con la RC. (2) Se han añadido las variables de personalidad ya que, los resultados para ambos sexos reflejan que los factores de personalidad son determinantes de las DP (véanse Tablas N1 y N2). La medida de personalidad empleada ha sido la quasi-ipsativa (QI5F-Tri) por presentar un mayor tamaño de muestra. (3) Finalmente, se han incluido las variables de DP, con la excepción de la variable dependencia-independencia de campo por tener un tamaño muestral más reducido. En el caso de las mujeres, este tercer paso ha sido repetido varias veces hasta depurar las variables que no han contribuido a la predicción de la RC. Los resultados así obtenidos son los recogidos en el Modelo 4.

Tabla N1.

Determinantes de las Disposiciones de Pensamiento (Hombres)

	Atención (β)	Impulsividad (β)	Intuición (β)	NFC (β)	Autocontrol (β)
WPT	.23**	.06*	-	.40**	.28**
EE	.83**	-.58**	-.21**	.68**	.46**
EX	1.64**	-.47**	-	1.46**	.98**
AP	1.24**	-.48**	-	1.47**	.52**
AM	1.14**	-.91**	-.35**	1.00**	.54**
C	1.81**	-1.20**	.53**	1.37**	1.45**
<i>F</i>	125.28	192.88	54.96	764.33	71.62
<i>R</i>	.82**	.87**	.56**	.96**	.73**
<i>R</i> ²	.67	.76	.30	.93	.53
<i>R</i> _{CV}	.81	.87	.54	.96	.72
<i>R</i> ² _{CV}	.66	.75	.29	.93	.52

Nota. WPT = Wonderlic personnel test; EE = estabilidad; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

p* < .05; *p* < .001

Tabla N2.

Determinantes de las Disposiciones de Pensamiento (Mujeres)

	Riesgo (β)	Impulsividad (β)	Intuición (β)	NFC (β)	Esfuerzo (β)
WPT	-	-	-	.41**	.12**
EE	-	-.37**	-.16**	.39**	.42**
EX	-	-.17**	-.17*	.74**	.83**
AP	.10**	-.29**	-.09	.93**	.41**
AM	-.14**	-.65**	-.39**	.42**	.29**
C	-.76**	-.86**	-.64**	.61**	1.17**
<i>F</i>	365.78	224.29	55.93	184.80	77.33
<i>R</i>	.78**	.78**	.53*	.78**	.62**
<i>R</i> ²	.60	.60	.27	.60	.38
<i>R</i> _{CV}	.77	.77	.51	.77	.61
<i>R</i> ² _{CV}	.60	.60	.26	.59	.37

Nota. WPT = Wonderlic personnel test; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

p* < .01; *p* < .001

Los análisis de regresión se han llevado a cabo mediante el programa SPSS y se han calculado, además del coeficiente de correlación múltiple (*R*) y el coeficiente de determinación múltiple (*R*²), el coeficiente de correlación múltiple cruzado de la población (*R*_{CV}) y el coeficiente de determinación múltiple cruzado de la población (*R*²_{CV}), mediante la fórmula de Browne (1975).

Las matrices de entrada utilizadas en los análisis de regresión se muestran en las Tablas N3 y N4 para hombres y mujeres, respectivamente. En ambos casos, se han empleado las correlaciones corregidas por error en la medida del predictor y del criterio y por restricción indirecta en el rango del predictor, según cada caso concreto.

Tabla N3.

Matriz de Correlaciones Corregidas Utilizada en el Análisis de Regresión para Determinar el Desempeño en el TRC-13 (Hombres)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. WPT	-												
2. Atención	.080	-											
3. Riesgo	.020	-.004	-										
4. NFC	.400	.589	.178	-									
5. Autocontrol	.062	.595	-.222	.385	-								
6. Impulsividad	.290	-.197	.612	.173	-.265	-							
7. Esfuerzo	.063	.794	-.144	.514	.770	-.238	-						
8. Intuición	.154	.212	.440	.437	-.072	.601	.082	-					
9. EE	.047	.089	.170	.022	.015	-.171	.028	-.081	-				
10. EX	.076	.022	.418	.094	-.082	.592	.023	.408	-.116	-			
11. AP	.212	.209	.351	.657	-.116	.151	.081	.261	-.112	-.214	-		
12. AM	-.098	-.086	-.013	-.120	-.241	-.346	.171	-.201	-.133	-.264	-.204	-	
13. C	-.259	.212	-.782	-.169	.542	-.566	.454	-.420	-.149	-.538	-.223	-.224	-
14. TRC-13	.690	.172	.276	.368	-.032	.076	.107	.087	.109	.090	.174	-.052	-.278

Nota. WPT = Wonderlic personnel test; NFC = need for cognition; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a).

Tabla N4.
Matriz de Correlaciones Corregidas Utilizada en el Análisis de Regresión para Determinar el Desempeño en el TRC-13 (Mujeres)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. WPT	-												
2. Atención	.244	-											
3. Riesgo	.134	-.049	-										
4. NFC	.389	.499	.346	-									
5. Autocontrol	-.154	.466	-.462	.138	-								
6. Impulsividad	.124	-.061	.645	.234	-.445	-							
7. Esfuerzo	-.044	.759	-.208	.382	.791	-.168	-						
8. Intuición	.118	.172	.470	.478	-.091	.736	.050	-					
9. EE	-.011	.161	.151	.010	.070	-.102	-.003	.021	-				
10. EX	-.129	-.080	.473	.048	-.111	.573	-.038	.335	-.033	-			
11. AP	.224	.077	.251	.624	-.281	.084	-.048	.156	-.221	-.174	-		
12. AM	-.204	-.121	-.020	-.135	-.102	-.412	-.225	-.215	.044	-.266	-.190	-	
13. C	-.071	.230	-.752	-.193	.516	-.490	.439	-.416	-.268	-.589	-.165	-.183	-
14. TRC-13	.747	.170	.100	.285	-.112	-.013	.004	.017	-.095	-.121	.211	-.002	.010

Nota. WPT = Wonderlic personnel test; NFC = need for cognition; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; TRC-13 = 3 ítems Frederick (2005) y 10 ítems Salgado (2014a).

Los resultados de los análisis se presentan en la Tabla N5, para la muestra de hombres y en la Tabla N6, para el grupo de mujeres. Como se puede observar, el modelo donde se incluyen conjuntamente las variables de CMG, personalidad y DP es el modelo que explicó mayor proporción de varianza en la muestra de hombres ($R^2_{CV} = .66$ modelo 3) y en la muestra de mujeres ($R^2_{CV} = .67$ modelo 4). No obstante, el incremento de varianza explicada respecto del modelo 2 (CMG y DP) y del modelo 1 (CMG) fue mayor en el grupo de hombres ($\Delta R^2_{CV} = .20$ y $\Delta R^2_{CV} = .18$, respectivamente) que en el de las mujeres ($\Delta R^2_{CV} = .04$ y $\Delta R^2_{CV} = .11$, respectivamente).

Tabla N5.

Análisis de Regresión Múltiple para Determinar el Desempeño en el TRC-13 de los Hombres

	Modelo 1 (β)	Modelo 2 (β)	Modelo 3 (β)
WPT	.69*	.66*	.62*
EE	-	.08	-.29*
EX	-	.03	-
AP	-	.04	-.64*
AM	-	.02	-.58*
C	-	-.07	-.49*
Autocontrol	-	-	-.58*
Impulsividad	-	-	-.68*
Atención	-	-	.16*
Esfuerzo	-	-	-
Intuición	-	-	-.17*
Riesgo	-	-	-
NFC	-	-	.71*
<i>F</i>	338.06	59.05	80.70
<i>R</i>	.69*	.70*	.83*
R^2	.48	.48	.68
R_{CV}	.69	.68	.81
R^2_{CV}	.48	.46	.66

Nota. WPT = wonderlic personnel test; NFC = need for cognition; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; *R* = coeficiente de correlación múltiple; R^2 = coeficiente de determinación múltiple; R_{CV} = coeficiente de correlación múltiple de la población cruzada; R^2_{CV} = coeficiente de determinación múltiple de la población cruzada.

* $p < .001$

Tabla N6.

Análisis de Regresión Múltiple para Determinar el Desempeño en el TRC-13 de las Mujeres

	Modelo 1 (β)	Modelo 2 (β)	Modelo 3 (β)	Modelo 4 (β)
WPT	.75***	.86***	.76*	.97***
EE	-	.10**	-.03	-
EX	-	.44***	-	.50***
AP	-	.28***	.18***	.43***
AM	-	.43***	.17***	.46***
C	-	.48***	.28***	.62***
Autocontrol	-	-	-.14**	-
Impulsividad	-	-	-.18**	-.20***
Atención	-	-	.05	-
Esfuerzo	-	-	-	.08*
Intuición	-	-	.13**	.23***
Riesgo	-	-	.21***	.28***
NFC	-	-	-.15**	-.38***
<i>F</i>	925.41	220.32	107.45	155.67
<i>R</i>	.75***	.80***	.79***	.83***
<i>R</i> ²	.56	.64	.62	.68
<i>R</i> _{CV}	.75	.80	.78	.82
<i>R</i> ² _{CV}	.56	.63	.61	.67

Nota. WPT = Wonderlic personnel test; EE = estabilidad emocional; EX = extroversión; AP = apertura a la experiencia; AM = amigabilidad; C = conciencia; *R* = coeficiente de correlación múltiple; *R*² = coeficiente de determinación múltiple; *R*_{CV} = coeficiente de correlación múltiple de la población cruzada; *R*²_{CV} = coeficiente de determinación múltiple de la población cruzada.

p* < .05; *p* < .01; ****p* < .001

En el caso de los hombres, aunque la CMG fue uno de los predictores más robusto de la RC (rango β = .62 a .69), el peso del coeficiente de regresión de la CMG disminuyó cuando se introdujeron en el modelo las variables de DP. Las variables NFC (β = .71), impulsividad (β = -.68) y AP (β = -.64) mostraron los pesos de regresión más altos. A estos les siguieron CMG (β = .62), autocontrol (β = -.58), AM (β = -.58), C (β = -.49), EE (β = -.29), intuición (β = -.17) y atención (β = .16). Aunque el factor de personalidad de EX no predijo directamente la RC en hombres si produce un efecto indirecto a través de las DP que determinan el desempeño en el TRC. Así pues, la Figura N1 representa gráficamente los efectos directos de las variables en la predicción de la RC de los hombres y las Figuras N2, N3, N4, N5 y N6 presentan los efectos indirectos o moderadores. Los índices de ajuste del modelo resultante de la muestra de hombres se recogen en la Tabla N7.

Del mismo modo, los resultados demostraron que la CMG fue el determinante más robusto de la RC de las mujeres, incluso cuando se controló por los efectos de las demás variables

disposicionales (rango $\beta = .75$ a $.97$). A esta le siguieron las dimensiones de personalidad C ($\beta = .62$), EX ($\beta = .50$), AM ($\beta = .46$) y AP ($\beta = .43$). Finalmente, las DP de NFC ($\beta = -.38$), Riesgo ($\beta = .28$), Intuición ($\beta = .23$), Impulsividad ($\beta = -.20$) y Esfuerzo ($\beta = .08$) también contribuyeron a explicar el desempeño en el TRC de las mujeres. Del mismo modo, aunque el factor de personalidad de EE no predijo de forma directa la RC de las mujeres, sí produce un efecto indirecto a través de algunas de las DP que determinan el desempeño en el test. Así, la Figura N7 refleja los efectos directos de las variables en la determinación de la RC y las Figuras N8, N9, N10, N11 y N12 representan gráficamente los efectos moderadores o indirectos sobre la RC. Los índices de ajuste del modelo se muestran en la Tabla N7.

Tabla N7.

Índices de Ajuste del Modelo de Predicción de la RC de Hombres y Mujeres

	χ^2	CFI	RMSEA	RMR
Hombres	205.68 ($p = .00$)	.91	.19	.06
Mujeres	1714.98 ($p = .00$)	.74	.38	.09

Nota. CFI = índice de bondad de ajuste comparativo; RMSEA = raíz del residuo cuadrático promedio de aproximación; RMR = raíz del residuo cuadrático promedio.

En resumen, los resultados mostraron que tanto variables cognitivas como de personalidad determinan el éxito en el TRC de hombres y mujeres y que, en conjunto, todas ellas explican la misma proporción de varianza en ambos grupos ($R^2_{CV} = .66$ en hombres y $R^2_{CV} = .67$ en mujeres). No obstante, las variables determinantes y la contribución de cada una de ellas parecen diferir en función del sexo. Así, mientras que la CMG se mostró como el determinante más relevante de la RC en mujeres ($\beta = .97$), en hombres fue el cuarto más robusto ($\beta = .62$). En la muestra de hombres, las variables de personalidad (NFC e impulsividad y el factor de AP) fueron las que más contribuyeron en la explicación de la RC ($\beta = .71$, $\beta = -.68$ y $\beta = -.64$, respectivamente). La DP de riesgo resultó ser un determinante independiente de la RC de las mujeres ($\beta = .28$), pero, sin embargo, en el caso de los hombres no lo fue. Del mismo modo, las DP de autocontrol y atención contribuyeron a explicar varianza de la RC de los hombres ($\beta = -.58$ y $\beta = .16$, respectivamente), pero no de las mujeres. Estos resultados, parecen apoyar la afirmación de Campitelli y Gerrans (2014) de que diferentes variables podrían explicar el desempeño en el TRC en hombres y mujeres.

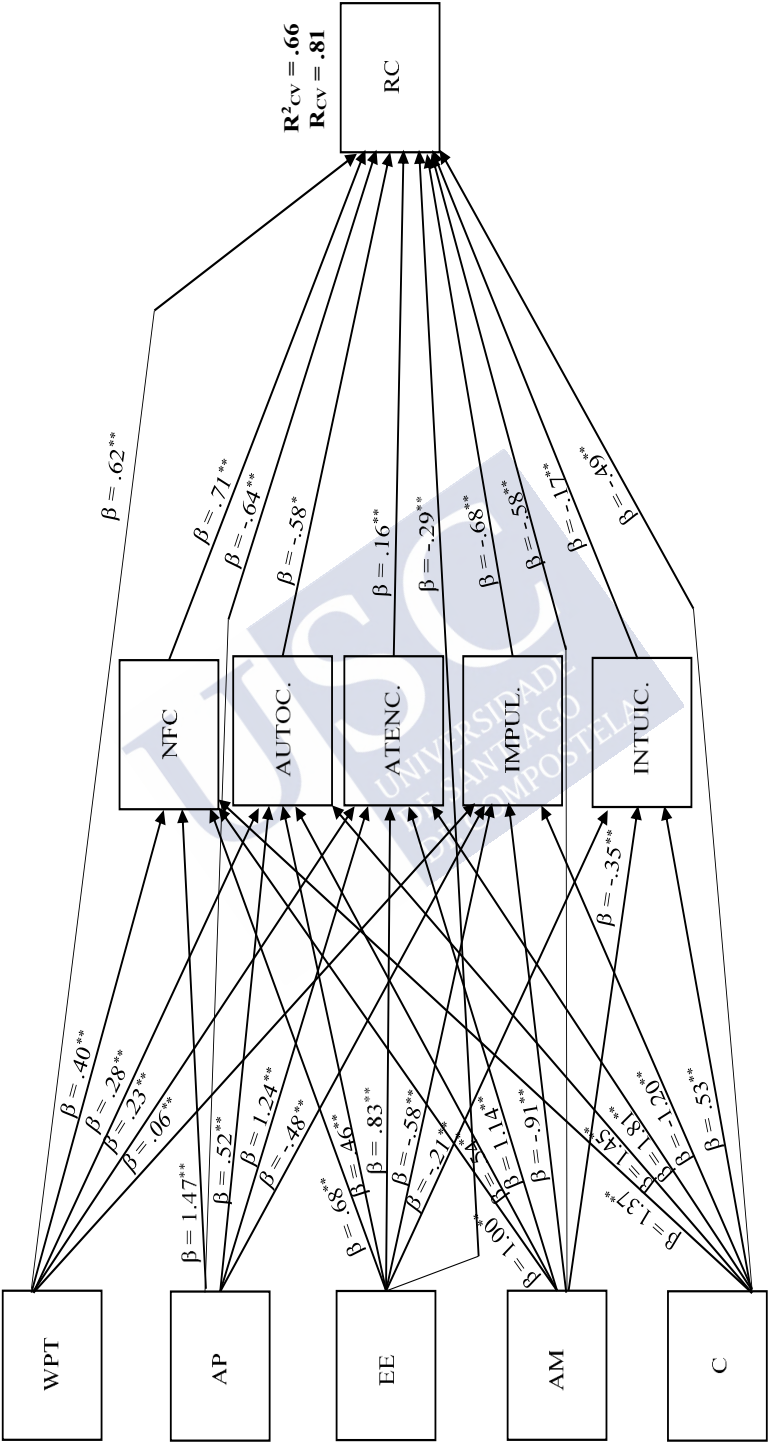


Figura N1. Efectos Directos de las Variables en la Predicción de la RC (Hombres)
Nota. $^* p < .05$; $^{**} p < .001$

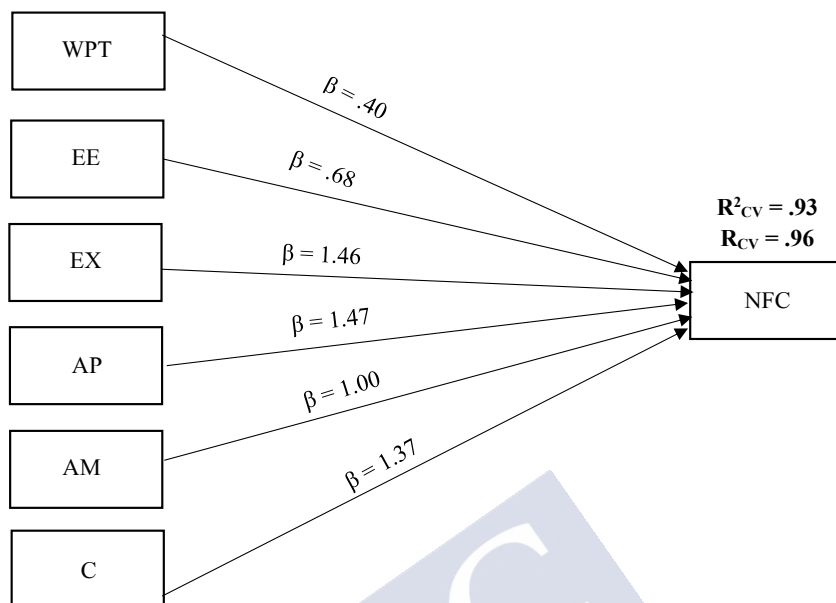


Figura N2. Efectos Indirectos en la Predicción de la RC a través de NFC (Hombres)

Nota. Todas las cargas factoriales son significativas a nivel $p < .001$

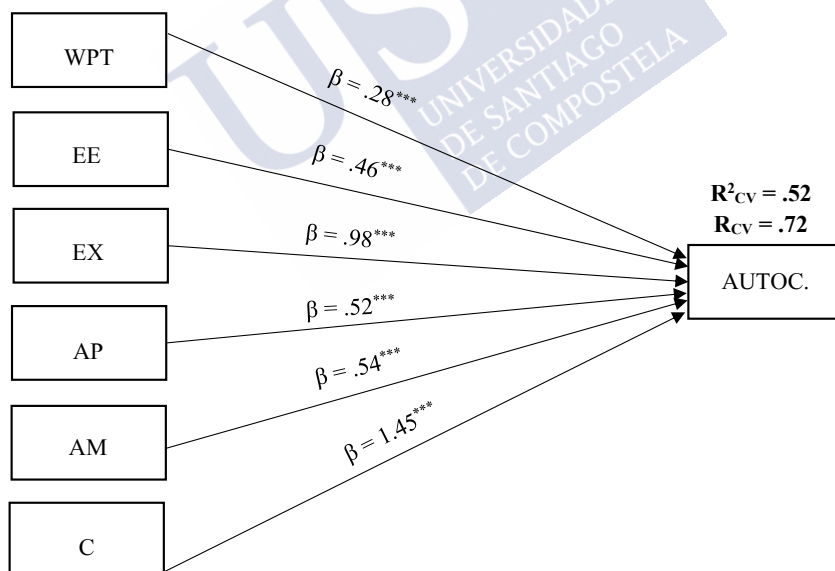


Figura N3. Efectos Indirectos en la Predicción de la RC a través de Autocontrol (Hombres)

Nota. Todas las cargas factoriales son significativas a nivel $p < .001$

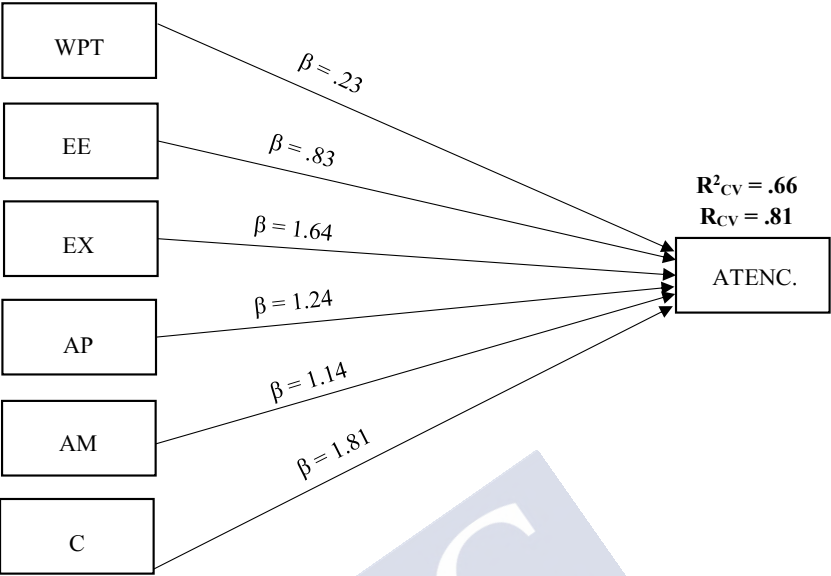


Figura N4. Efectos Indirectos en la Predicción de la RC a través de Atención (Hombres)
Nota. Todas las cargas factoriales son significativas a nivel $p < .001$

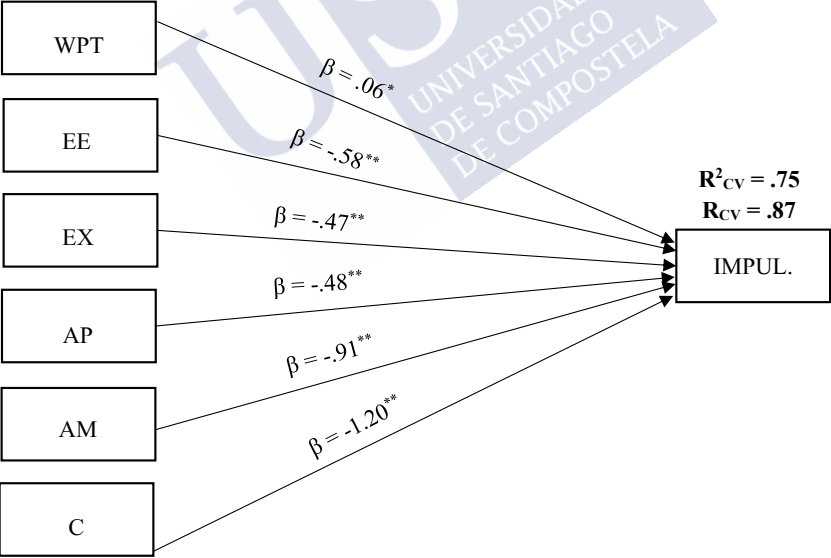


Figura N5. Efectos Indirectos en la Predicción de la RC a través de Impulsividad (Hombres)
Nota. $^*p < .05$; $^{**}p < .001$

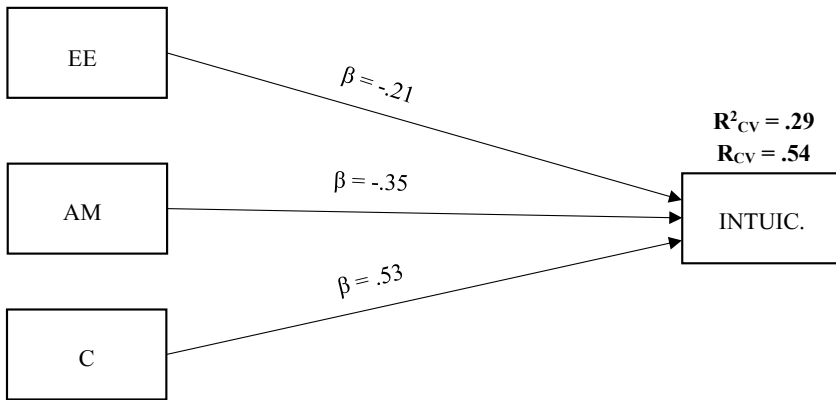


Figura N6. *Efectos Indirectos en la Predicción de la RC a través de Intuición (Hombres)*

Nota. Todas las cargas factoriales son significativas a nivel $p < .001$



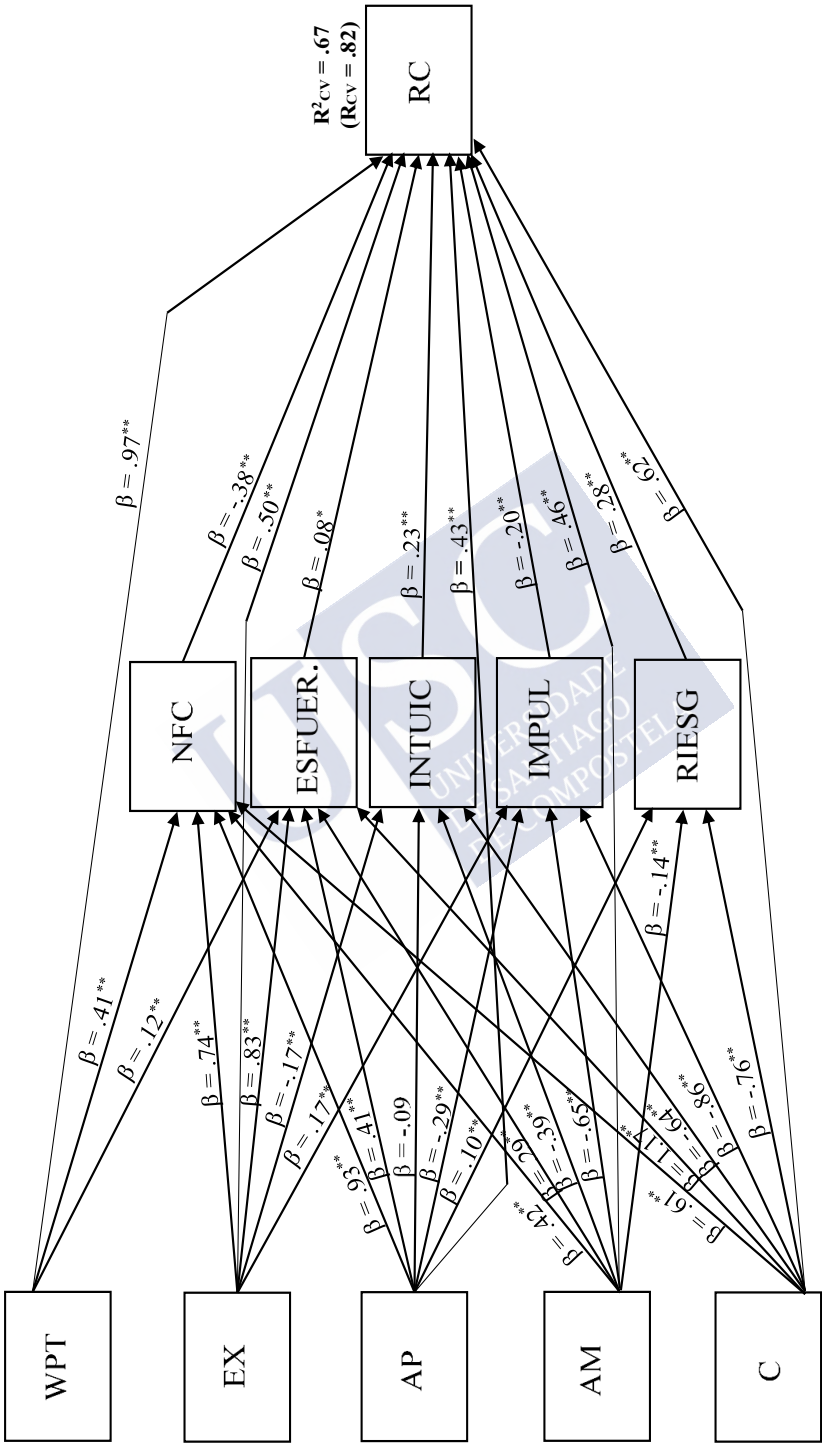


Figura N7. Efectos Directos de las Variables en la Predicción de la RC (Mujeres)

Nota. * $p < .05$; ** $p < .001$

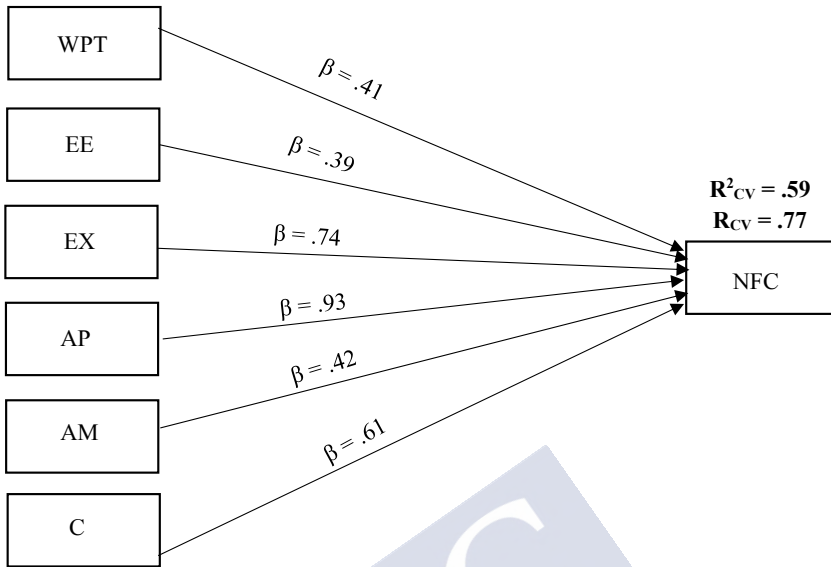


Figura N8. *Efectos Indirectos en la Predicción de la RC a través de NFC (Mujeres)*

Nota. Todas las cargas factoriales son significativas a nivel $p < .001$

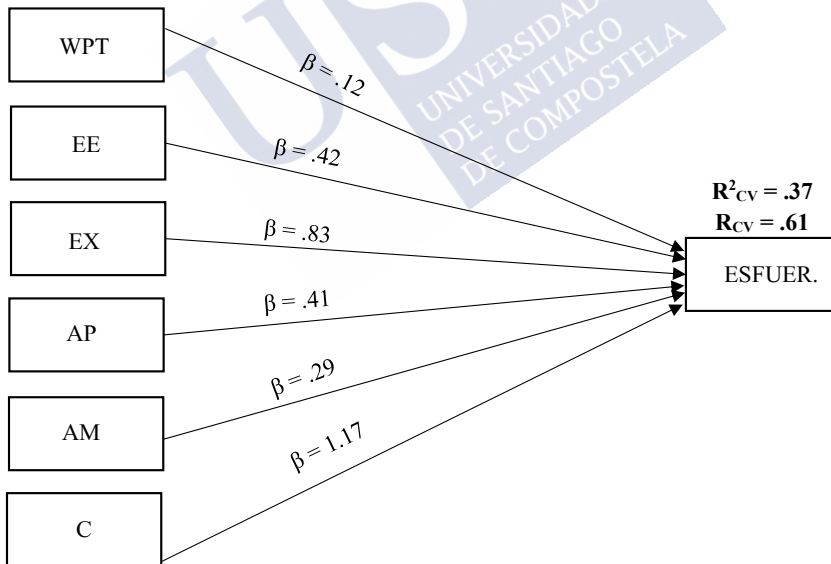


Figura N9. *Efectos indirectos en la predicción de la RC a través de Esfuerzo (mujeres)*

Nota. Todas las cargas factoriales son significativas a nivel $p < .001$

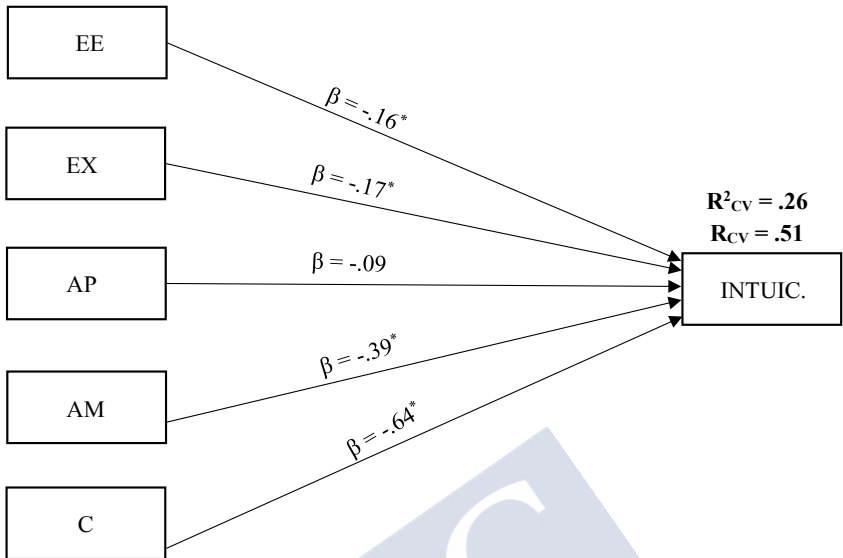


Figura N10. *Efectos Indirectos en la Predicción de la RC a través de Intuición (Mujeres)*
Nota. * $p < .001$

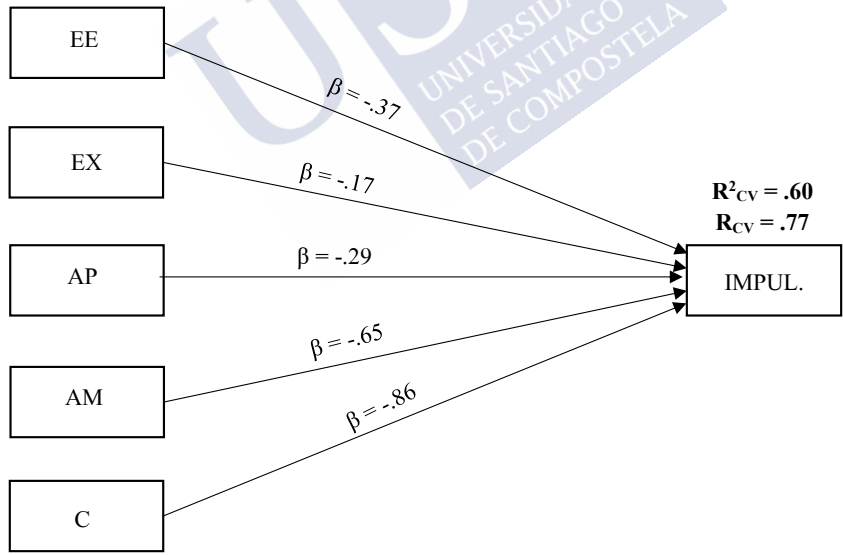


Figura N11. *Efectos Indirectos en la Predicción de la RC a través de Impulsividad (Mujeres)*
Nota. Todas las cargas factoriales son significativas a nivel $p < .001$

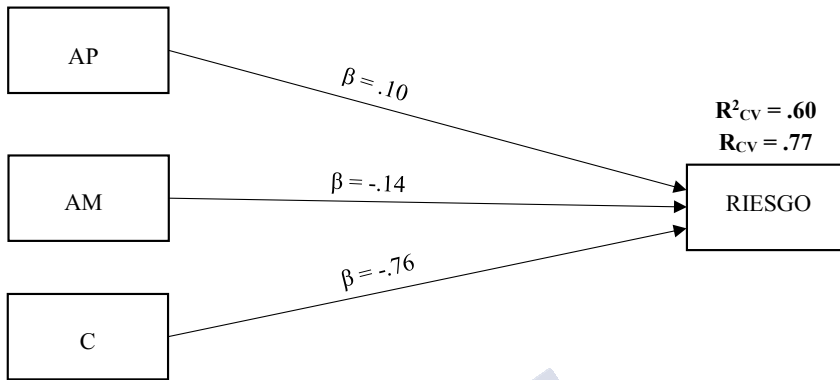


Figura N12. *Efectos Indirectos en la Predicción de la RC a través de Riesgo (Mujeres)*

Nota. Todas las cargas factoriales son significativas a nivel $p < .001$



APÉNDICE Ñ.

Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Relación de la CMG y la RC Excluyendo las Numeracy Scales

Referencia	Tipo de TRC	Tipo de CC	N	r_{xy}	r_{xx}	Estudio
Bergman et al. (2010)	Original	CMG	116	.64	-	
Brañas-Garza et al. (2012)	Original	CMG	191	.29	-	
Carpenter et al. (2013)	Original	SAT/ACT	422	.35	-	
Cokely et al. (2012)	Original	Compuesto	300	.34	.77	
Corgnet et al. (2016)	TRC-7	CMG	150	.26	-	
Corgnet, Espín, Hernán-Gonzalez (2015)	TRC-7	CMG	150	.43	-	
Corgnet, Hernán-González et al. (2015)	Original	SAT	264	.35	-	
Del Missier et al. (2011)	Original	CMG	206	.22	.84	
Fehr y Huck (2015)	Original	CMG	240	.47	-	
Finucane y Gullion (2010)	TRC-6	Compuesto	108	.41	-	
Frederick (2005)	Original	CMG	921	.43	-	
Gino y Ariely (2012)	Original	CC	97	.30	-	
Goldwater et al. (2018)	Original	CMG	86	.55	-	
Hanaki et al. (2016)	Original	CMG	200	.31	-	
Insler et al. (2015)	Original	SAT	364	.39	-	
Jackson et al. (2016)	TRC-7	Compuesto	250	.57	.64	
Liberali et al. (2012)	Original	CMG	190	.31	.77	
Lindeman y Svedholm-Häkkinen (2016)	Original	Compuesto	258	.40	-	
Moritz et al. (2014)	Original	CMG	147	.44	-	
Morsanyi et al. (2012)	Original	Compuesto	123	.42	-	
Morsanyi et al. (2017)	TRC-6	Compuesto	69	.46	.87	
Narayanan y Moritz (2015)	Original	CMG	60	.43	-	
Obrecht et al. (2009)	Original	SAT	417	.45	-	
Pennycook et al. (2012)	Original	CC	231	.37	-	Est. 1
	Original	CC	267	.24	-	Est. 2
Pennycook et al. (2014a)	Original	CC	196	.40	-	
Pennycook, Cheyne, Barr et al. (2015)	TRC-7	CC	279	.41	.65	
Poore et al. (2014)	Original	Compuesto	431	.51	-	
^a Primi et al. (2015)	Original	CMG	201	.32	-	
	TRC-6	CMG	201	.39	-	

Continúa

Continuación

Referencia	Tipo de TRC	Tipo de CC	<i>N</i>	r_{xy}	r_{xx}	Estudio
Razmyar y Reeve (2013)	Original	CMG	150	.53	-	
Ruffle y Tobol (2017)	Original	Compuesto	427	.70	-	
Shenhav et al. (2011)	Original	Compuesto	302	.38	-	
Shenhav et al. (2017)	Original	Compuesto	335	.38	-	
Sinayev y Peters (2015)	TRC-5	CMG	1,135	.24	-	
Teovanović et al. (2015)	Original	Compuesto	243	.26	.88	
Thompson et al. (2013)	Original	SAT	150	.34	-	Exp. 3a
	Original	CMG	142	.43	-	Exp. 3b
^a Thomson y Oppenheimer (2016)	Original	SAT	143	.59	-	
	TRC-7	SAT	143	.59	-	
Toplak et al. (2011)	Original	Compuesto	346	.32	-	
^a Toplak et al. (2014)	Original	CMG	160	.48	-	
	TRC-7	CMG	160	.50	-	
Trippas et al. (2015)	Original	CC	182	.39	-	
Ventis (2015)	Original	SAT	148	.41	-	
Welsh et al. (2013)	Original	CC	102	.44	-	
West et al. (2012)	Original	SAT	482	.45	-	
^a Estudio 3 de la tesis	Original	CMG	1,109	.41	.88	
	TRC-13	CMG	1,109	.48	.88	

Nota. Tipo de CC = medida de capacidad cognitiva empleada; *N* = tamaño de la muestra; r_{xy} = tamaño del efecto observado; r_{xx} = fiabilidad por consistencia interna de la medida de capacidad cognitiva; Original = TRC de Frederick (2005); CMG = capacidad mental general; SAT/ACT = pruebas cognitivas estandarizadas de admisión académica; Compuesto = se calculó un compuesto con todas las medidas de capacidad cognitiva reportadas para esa muestra; CC = medida de capacidad cognitiva específica (medida imperfecta del constructo).

^aLos tamaños del efecto corresponden a la misma muestra, pero usando un TRC diferente.

APÉNDICE O.

Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Relación de la CMG y la RC Incluyendo las Numeracy Scales

Referencia	Tipo de TRC	Tipo de CC	N	r_{xy}	r_{xx}	Estudio
Bergman et al. (2010)	Original	CMG	116	.64	-	
Brañas-Garza et al. (2012)	Original	CMG	191	.29	-	
Campitelli y Gerrans (2014)	Original	CC	2,019	.43	.51	
Carpenter et al. (2013)	Original	SAT/ACT	422	.35	-	
Cokely y Kelley (2009)	Original	CC	69	.31	-	
Cokely et al. (2012)	Original	Compuesto	300	.41	.73	
Corgnet et al. (2016)	TRC-7	CMG	150	.26	-	
Corgnet, Espín, Hernán-Gonzalez (2015)	TRC-7	CMG	150	.43	-	
Corgnet, Hernán-González et al. (2015)	Original	SAT	264	.35	-	
Del Missier et al. (2011)	Original	Compuesto	206	.35	.45	
Drummond y Fischhoff (2017)	Original	CC	395	.46	.55	Est. 1c
	Original	CC	270	.49	.52	Est. 2
Fehr y Huck (2015)	Original	CMG	240	.47	-	
Finucane y Gullion (2010)	TRC-6	CC	608	.53	.62	
Frederick (2005)	Original	CMG	921	.43	-	
Gino y Ariely (2012)	Original	CC	97	.30	-	
Goldwater et al. (2018)	Original	CMG	86	.55	-	
Graffeo et al. (2015)	Original	CC	35	.29	-	
Hanaki et al. (2016)	Original	CMG	200	.31	-	
Insler et al. (2015)	Original	SAT	364	.39	-	
Jackson et al. (2016)	TRC-7	Compuesto	250	.57	.64	
Liberali et al. (2012)	Original	CC	259	.51	.71	Est. 1
	Original	Compuesto	190	.44	.44	Est. 2
Lindeman y Svedholm-Häkkinen (2016)	Original	Compuesto	258	.40	-	
Mata (2016)	Original	CC	104	.47	-	
Moritz et al. (2014)	Original	CMG	147	.44	-	
Morsanyi et al. (2012)	Original	Compuesto	123	.42	-	
Morsanyi et al. (2017)	TRC-6	Compuesto	69	.46	.87	
Narayanan y Moritz (2015)	Original	CMG	60	.43	-	
Obrecht et al. (2009)	Original	Compuesto	417	.50	-	

Continúa

Continuación

Referencia	Tipo de TRC	Tipo de CC	N	r_{xy}	r_{xx}	Estudio
Pennycook et al. (2012)	Original	CC	231	.37	-	Est. 1
	Original	CC	267	.24	-	Est. 2
Pennycook et al. (2014a)	Original	CC	196	.40	-	
Pennycook, Cheyne, Barr et al. (2015)	TRC-7	Compuesto	279	.49	.46	
Poore et al. (2014)	Original	Compuesto	431	.51	-	
^a Primi et al. (2015)	Original	Compuesto	201	.37	.71	
	TRC-6	Compuesto	201	.41	.71	
Primi et al. (2018)	TRC-6	CC	181	.42	.64	
Razmyar y Reeve (2013)	Original	CMG	150	.53	-	
Ruffle y Tobol (2017)	Original	Compuesto	427	.70	-	
Shenhav et al. (2011)	Original	Compuesto	302	.38	-	
Shenhav et al. (2017)	Original	Compuesto	335	.38	-	
Sinayev y Peters (2015)	TRC-5	CC	1,409	.57	.67	Est. 1
	TRC-5	Compuesto	1,135	.24	-	Est. 2
Sirota y Juanchich (2011)	Original	CC	94	.52	.74	
Skagerlund et al. (2018)	Original	CC	2,063	.61	.58	
^a Szaszi et al. (2017)	Original	CC	178	.49	-	
	TRC-6	CC	178	.29	-	
Teovanović et al. (2015)	Original	Compuesto	243	.26	.88	
Thompson et al. (2013)	Original	SAT	150	.34	-	Exp. 3a
	Original	CMG	142	.43	-	Exp. 3b
^a Thomson y Oppenheimer (2016)	Original	CC	136	.66	.75	
	TRC-7	CC	136	.63	.75	
Toplak et al. (2011)	Original	Compuesto	346	.32	-	
^a Toplak et al. (2014)	Original	CMG	160	.48	-	
	TRC-7	CMG	160	.50	-	
Trippas et al. (2015)	Original	Compuesto	182	.46	-	
Ventis (2015)	Original	SAT	148	.41	-	
Weller et al. (2013)	Original	CC	1,97	.49	.76	
Welsh et al. (2013)	Original	CC	102	.44	-	
West et al. (2012)	Original	SAT	482	.45	-	
^a Estudio 3 de la tesis	Original	CMG	1,109	.41	.88	
	TRC-13	CMG	1,109	.48	.88	

Nota. Tipo de CC = medida de capacidad cognitiva empleada; N = tamaño de la muestra; r_{xy} = tamaño del efecto observado; r_{xx} = fiabilidad por consistencia interna de la medida de capacidad cognitiva; Original = TRC de Frederick (2005); CMG = capacidad mental general; SAT/ACT = pruebas cognitivas estandarizadas de admisión académica; Compuesto = se calculó un compuesto con todas las medidas de capacidad cognitiva reportadas para esa muestra; CC = medida de capacidad cognitiva específica (medida imperfecta del constructo).

^aLos tamaños del efecto corresponden a la misma muestra, pero usando un TRC diferente.

APÉNDICE P.

Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Relación entre las Capacidades Cognitivas Específicas y la RC

Estudio Primario	Tipo de TRC	Tipo de CC	N	r_{xy}	r_{xx}	Estudio
Campitelli y Gerrans (2014)	Original	<i>Numeracy</i>	2,019	.43	.51	
Cokely et al. (2012)	Original	Compuesto <i>Numeracy</i>	300	.56	.67	
	Original	CV	300	.28	.80	
	Original	CMO	300	.26	-	
Cokely y Kelley (2009)	Original	CMO	69	.31	-	
Del Missier et al. (2011)	Original	<i>Numeracy</i>	206	.34	.73	
Drummond y Fischhoff (2017)	Original	<i>Numeracy</i>	395	.46	.55	Est. 1c
	Original	<i>Numeracy</i>	270	.49	.52	Est. 2
Finucane y Gullion (2010)	TRC-6	<i>Numeracy</i>	608	.53	.62	
	TRC-6	CV	135	.37	-	
	TRC-6	CMO	129	.36	-	
Gino y Ariely (2012)	Original	CV	97	.30	-	
Goldwater et al. (2018)	Original	CMO	60	.25	-	
Gómez-Chacón et al. (2014)	TRC-5	CMO	56	.17	-	
Graffeo et al. (2015)	Original	<i>Numeracy</i>	35	.29	-	
Jackson et al. (2016)	TRC-7	CV	250	.38	.74	
Liberali et al. (2012)	Original	<i>Numeracy</i>	259	.51	.71	Est. 1
	Original	<i>Numeracy</i>	190	.39	.82	Est. 2
Lindeman y Svedholm-Häkkinen (2016)	Original	CME	258	.40	-	
Mata (2016)	Original	<i>Numeracy</i>	104	.47	-	
Morsanyi et al. (2012)	Original	CME	123	.42	-	
Morsanyi et al. (2017)	TRC6	CN	69	.40	.72	
Pennycook et al. (2012)	Original	CV	231	.37	-	Est. 1
	Original	CV	267	.24	-	Est. 2
Pennycook et al. (2014a)	Original	CV	196	.40	-	
Pennycook, Cheyne, Barr et al. (2015)	TRC-7	CV	279	.41	.65	
	TRC-7	<i>Numeracy</i>	279	.38	.47	
Poore et al. (2014)	Original	CN	431	.50	-	
^a Primi et al. (2015)	Original	<i>Numeracy</i>	201	.41	.71	
	Original	CN	59	.28	-	

Continúa

Continuación

Estudio Primario	Tipo de TRC	Tipo de CC	N	r_{xy}	r_{xx}	Estudio
*Primi et al. (2015)	TRC-6	<i>Numeracy</i>	201	.44	.71	
	TRC-6	CN	59	.42	-	
Primi et al. (2018)	TRC-6	<i>Numeracy</i>	181	.42	.64	
Sevenants et al. (2011)	Original	CMO	91	.23	-	Exp. 1
	Original	CMO	77	.39	-	Exp. 2
Shenhav et al. (2011)	Original	CV	307	.27	-	
Shenhav et al. (2017)	Original	CV	335	.29	-	
Sinayev y Peters (2015)	TRC-5	<i>Numeracy</i>	1,409	.57	.67	Est. 1
	TRC-5	<i>Numeracy</i>	1,459	.51	.58	Est. 2
Sirota y Juanchich (2011)	Original	<i>Numeracy</i>	94	.52	.74	
Skagerlund et al. (2018)	Original	<i>Numeracy</i>	2,063	.61	.58	
Stuppel et al. (2013)	Original	CMO	65	.53	-	Exp. 1
	Original	CMO	49	.50	-	Exp. 2
*Szasz et al. (2017)	Original	<i>Numeracy</i>	178	.49	-	
	TRC-6	<i>Numeracy</i>	178	.29	-	
Teovanović et al. (2015)	Original	CMO	243	.24	.79	
	Original	CME	243	.35	.79	
	Original	CV	243	.13	.71	
*Thomson y Oppenheimer (2016)	Original	<i>Numeracy</i>	136	.58	.75	
	TRC-7	<i>Numeracy</i>	136	.52	.75	
Toplak et al. (2011)	Original	CV	346	.17	-	
	Original	CMO	346	.33	-	
Trippas et al. (2015)	Original	CV	182	.39	-	
	Original	<i>Numeracy</i>	182	.30	-	
Weller et al. (2013)	Original	<i>Numeracy</i>	1,97	.49	.76	
Welsh et al. (2013)	Original	CN	102	.44	-	

Nota. Tipo de CC = medida de capacidad cognitiva específica empleada; N = tamaño de la muestra; r_{xy} = tamaño del efecto observado; r_{xx} = fiabilidad por consistencia interna de la medida de capacidad cognitiva; Original = TRC de Frederick (2005); CN = capacidad numérica; CV = capacidad verbal; CME = capacidad mecánico-espacial; CMO = capacidad de memoria operativa; *Numeracy* = numeracy scale.

*Los tamaños del efecto corresponden a la misma muestra, pero usando un TRC diferente.

APÉNDICE Q.

Datos de los Estudios Primarios Incluidos en el Meta-Análisis de la Validez de la RC para Predecir el Desempeño Académico

Estudio Primario	Tipo de TRC	<i>N</i>	<i>r_{xy}</i>	Estudio
Corgnet, Hernán-Gonzalez et al. (2015)	Original	264	.04	
Gómez-Chacón et al. (2014)	TRC-5	56	.25	
Insler et al. (2015)	Original	364	.31	
Lindeman y Svedholm-Häkkinen (2016)	Original	258	.34	
Lohse (2016)	Original	284	.27	
Morsanyi et al. (2014)	Original	328	.45	Est. 1
	Original	184	.29	Est. 2
^a Primi et al. (2017)	Original	124	.17	
	TRC-6	124	.22	
^a Thomson y Oppenheimer (2016)	Original	143	.33	
	TRC-7	143	.37	
^a Toplak et al. (2014)	Original	160	.23	
	TRC-7	160	.25	
^a Estudio 3 de la tesis	Original	898	.16	
	TRC-13	898	.15	

Nota. *N* = tamaño de la muestra; *r_{xy}* = tamaño del efecto observado; Original = TRC de Frederick (2005).

^aLos tamaños del efecto corresponden a la misma muestra, pero usando un TRC diferente.



APÉNDICE R.

Datos de los Estudios Primarios Incluidos en el Meta-Análisis sobre Diferencias entre Hombres y Mujeres en RC

Estudios Primarios	Tipo de TRC	N	d	Estudio
Aczel et al. (2015)	Original	864	.256	
Aktas et al. (2017)	Original	267	.124	Est. 1
	Original	239	.345	Est. 2
	Original	382	-.165	Est. 3
Albaity et al. (2014)	Original	880	.486	
Alós-Ferrer et al. (2016)	Original	155	.476	
Alós-Ferrer y Hügelschäfer (2016)	Original	416	.827	Exp. 1
	Original	111	.528	Exp. 2
	Original	364	.335	Exp. Clase
Baron et al. (2015)	Original	103	.242	
Böckenholt (2012)	Original	579	.162	
Bosch-Doménech et al. (2014)	Original	623	.394	
Brosnan et al. (2014)	Original	68	.517	
Browne et al. (2014)	Original	1,137	.538	
Campitelli y Gerrans (2014)	Original	2,019	.336	
Campitelli y Labollita (2010)	Original	155	.220	
Corgnet, Espín y Hernán-Gonzalez et al. (2015)	TRC-7	150	.603	Est. 1
	TRC-7	158	.673	Est. 2
Corgnet et al. (2016)	TRC-7	150	.549	
Drummond y Fischhoff (2017)	Original	395	.383	Est. 1a
	Original	393	.324	Est. 1c
	Original	268	.307	Est. 2
Duttie y Inukai (2015)	Original	66	1.155	
Finucane y Gullion (2010)	Original	608	.408	
Frederick (2005)	Original	3,428	.112	
Grossman et al. (2014)	TRC-4	224	.346	
Guthrie et al. (2007)	Original	241	.132	
Kiss et al. (2016)	Original	900	.343	
Lohse (2016)	Original	284	.359	
Narayanan y Moritz (2015)	Original	96	.062	
Obrecht et al. (2009)	Original	417	.606	
Pennycook et al. (2012)	Original	223	.324	Est. 1
	Original	267	.473	Est. 2
Pennycook, Cheyne, Koehler et al. (2015)	Original	369	.510	
Ponti y Carbone (2009)	Original	48	.873	
Ponti et al. (2014)	Original	192	.512	

Continúa

Continuación

Estudios Primarios	Tipo de TRC	<i>N</i>	<i>d</i>	Estudio
^a Primi et al. (2015)	Original	939	.340	
	TRC-6	908	.320	
Primi et al. (2018)	Original	281	.719	Est. 1 ^b
	Original	282	.624	
	Original	282	.624	
	Original	282	.529	
	Original	282	.529	
	Original	282	.529	
	Original	282	.434	
	Original	282	.434	
	Original	281	.339	
	TRC-6	181	.699	Est. 2
Razmyar y Reeve (2013)	Original	150	.606	
Royzman et al. (2014)	Original	527	.372	
Royzman et al. (2015)	TRC-6	548	.335	
Schulze y Newell (2015)	Original	58	.817	Exp. 2a
	Original	58	.829	Exp. 2b
Sinayev y Peters (2015)	TRC-5	1,46	.473	
Skagerlund et al. (2018)	Original	2,058	.432	
^a Thomson y Oppenheimer (2016)	Original	129	.813	
	TRC-verbal	131	.248	
^a Toplak et al. (2014)	Original	158	.637	
	TRC-4	158	.652	
Toplak et al. (2017)	TRC-11	232	.473	
Ventis (2015)	Original	146	.449	
Welsh et al. (2013)	Original	102	.516	
Willard y Norenzayan (2017)	Original	1,006	.306	
Yilmaz y Saribay (2016)	Original	337	.293	Est. 1
	Original	691	.600	Est. 2
	Original	126	.128	Est. 3a
	Original	86	.473	Est. 3b
^a Yilmaz y Saribay (2017)	Original	395	.366	
	TRC-verbal	395	-.020	
^a Estudio 3 de la tesis	Original	1,367	.452	
	TRC-13	1,367	.625	

Nota. *N* = tamaño de la muestra; *d* = tamaño del efecto observado; Original = TRC de Frederick (2005).

^aLos tamaños del efecto corresponden a la misma muestra, pero usando un TRC diferente.

^bLos tamaños del efecto se obtuvieron a partir de la distribución meta-analítica calculada por los autores.